frontiers
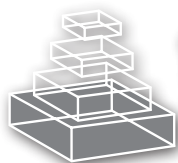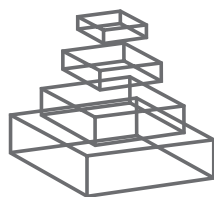# RESEARCH TOPICS

LANGUAGE BY MOUTH AND BY HAND

Topic Editors
Iris Berent and Susan Goldin-Meadow

ABOUT FRONTIERS

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

FRONTIERS JOURNAL SERIES

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing.

All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

DEDICATION TO QUALITY

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.

Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view.

By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

WHAT ARE FRONTIERS RESEARCH TOPICS?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area!

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: researchtopics@frontiersin.org

LANGUAGE BY MOUTH AND BY HAND

Topic Editors:
Iris Berent, Northeastern University, USA
Susan Goldin-Meadow, University of Chicago, USA

Source: lips http://www.morguefile.com/archive/display/537947
sign language http://pixabay.com/en/good-approval-green-light-okay-422550/
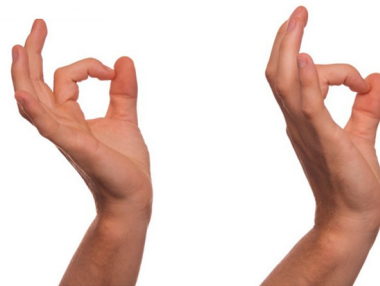
While most natural languages rely on speech, humans can spontaneously generate comparable linguistic systems that utilize manual gestures. This collection of papers examines the interaction between natural language and its phonetic vessels—human speech or manual gestures. We seek to identify what linguistic aspects are invariant across signed and spoken languages, and determine how the choice of the phonetic vessel shapes language structure, its processing and its neural implementation. We welcome rigorous empirical studies from a wide variety of perspectives, ranging from behavioral studies to brain analyses, diverse ages (from infants to adults), and multiple languages—both conventional and emerging home signs and sign languages.

Table of Contents

Language by mouth and by hand

Iris Berent[1]* and Susan Goldin-Meadow[2]

[1] Phonology and Reading Lab, Department of Psychology, Northeastern University, Boston, MA, USA
[2] Goldin-Meadow Laboratory, Department of Psychology, University of Chicago, Chicago, IL, USA
*Correspondence: i.berent@neu.edu

What is the basis of the human capacity for language: Is language shaped only by sensorimotor constraints and experience, or are some aspects of language universal, abstract, and potentially amodal? The set of papers assembled in this collection represent state of the art research on this age-old set of questions.

To gauge the universality of language structure and its abstraction, the first group of papers examines the grammatical organization of mature languages across modalities. The papers by Baus et al. (2014) and Guellaï et al. (2014) suggest that, despite marked differences in modality, the phonology of signed and spoken languages share aspects of design. Specifically, Baus and colleagues demonstrate that syllable-like units are extracted by signers automatically even when the task does not demand it. Using a similar interference paradigm, Guellaï and colleagues show that speakers (of Italian) automatically extract prosodic structure and use manual gestures to help them do it; the cues to prosody that are found in co-speech gesture play a role in disambiguating the syntactic structure of the speech they accompany. The typological survey described in Napoli and Sutton-Spence (2014) extends the study of grammatical universals to syntax. Like spoken languages, sign languages overwhelmingly favor subject-first structures (i.e., SOV and SVO). Unlike spoken languages, however, sign languages show a strong preference for the SOV over the SVO order. This aspect of grammatical organization may thus be influenced by modality, although the fact that signed and spoken languages differ not only with respect to modality but also with respect to age (i.e., spoken languages are older than sign languages) makes it difficult to pinpoint the source of this difference.

Further insights into grammar and its origins are presented in papers on the genesis of sign languages in Deaf communities and in individual homesigners (deaf individuals who have not been exposed to an established sign language and who use their own homemade gestures to communicate with the hearing individuals in their worlds). Given the poverty of linguistic input available to these individuals, and the fact that the manual modality affords iconic depiction, we might expect emerging sign languages to be overwhelmingly iconic. But the role of iconicity is actually far more constrained and nuanced than one might have presumed.

Considering homesigns, Coppola and Brentari (2014) find that the spontaneous emergence of morphophonology in an individual homesigning child mirrors the organization of mature sign languages (i.e., greater finger complexity in Object-handshapes than in Handling-handshapes). But remarkably, this abstract grammatical organization emerges *prior* to the arguably more iconic organization of morphosyntax (i.e., associating Object-handshapes with no-agent events and Handling-handshape with agent events). Moving to another example, this time a sign language that is growing up in Nicaragua, Kocab et al. (2015) find that, contrary to naïve expectations, signers do not immediately rely on iconic spatial devices to mark referential shifts, but rely instead on abstract lexical markers. Further glimpses into the spontaneous emergence of abstract syntactic organization can be found in Kastner et al. (2014), who document how prosody is used to mark the kernels of syntactic embedding in Kafr Qasem Sign Language, a sign language emerging in Israel.

The possibility that signed and spoken languages might both rely on abstract grammatical organization brings the ongoing debate between algebraic (symbolic, rule-based) vs. associationist accounts of spoken language into the domain of sign language—what computational mechanisms are used by signers to support linguistic productivity? The papers by Caselli and Cohen-Goldberg (2014), on one hand, and Berent et al. (2014), on the other hand, suggest that a full account of sign language computation (like spoken language computation) requires both systems, hence, "words and rules (Pinker, 1999)." Considering first the evidence for associations, Caselli and Cohen-Goldberg trace lexical competition in sign language to the same set of dynamic associative principles proposed for spoken languages. Nonetheless, Berent et al. find that signers can extend certain phonological generalizations across the board in a rule-governed way—even to novel signs with features that are unattested in their language. Building on past computational work, Berent et al. suggest that generalizations of this sort are the hallmark of powerful algebraic rules that support the capacity for discrete infinity in the manual modality.

Our review has so far highlighted commonalities across different language modalities and different levels of experience. But the effects of modality and experience are undeniable and significant—the papers by Supalla et al. (2014) and Emmorey et al. (2014) underscore some of these effects. Considering first experience, Supalla and colleagues find that language experience shapes language fluency, which, in turn, shapes the quality of signers' working-memory storage—fluent signers retain global semantic structure, less fluent signers focus on lexical detail and linear order. Considering language modality, Emmorey and colleagues find that, even though signed and spoken languages share neural substrates, sign language comprehension and production engages a unique network of sensorimotor regions that are

directly linked to the visual/manual channel; sign comprehension uniquely suppresses visual occipital activity, whereas sign production engages parietal regions involved in manual motor simulation.

The final four papers in this volume consider the development of sign languages and their evolution. Morgan (2014) argues that, across modalities, combinatorial structure emerges gradually out of a system that is initially holistic. Lillo-Martin et al. (2014) investigate the developmental of linguistic communication in bimodal bilingual children. Although these children are clearly sensitive to the language of their interlocutors and they modulate their language choice accordingly, the findings nonetheless reveal an overwhelming preference for speech over sign. In contrast, when adult speakers are engaged in a communication game, Fay et al. (2014) find a strong advantage for gestures over speech (alone, or even in combination with gesture)—a finding that the authors attribute to the affordance of the manual modality for iconicity. The gesture advantage in adult speakers does not speak directly to language evolution in humans, but the results are in line with the possibility that proto-language was gestural. How could such a gestural system give rise to the evolution of spoken language? This question is addressed by Woll (2014), who suggests that echo-phonology might provide the missing link.

REFERENCES

Baus, C., Gutiérrez, E., and Carreiras, M. (2014). The role of syllables in sign language production. *Front. Psychol.* 5:1254. doi: 10.3389/fpsyg.2014.01254

Berent, I., Dupuis, A., and Brentari, D. (2014). Phonological reduplication in sign language: rules rule. *Front. Lang. Sci.* 5:560. doi: 10.3389/fpsyg.2014.00560

Caselli, N. K., and Cohen-Goldberg, A. M. (2014). Lexical access in sign language: a computational model. *Front. Psychol.* 5:428. doi: 10.3389/fpsyg.2014.00428

Coppola, M., and Brentari, D. (2014). From iconic handshapes to grammatical contrasts: longitudinal evidence from a child homesigner. *Front. Psychol.* 5:830. doi: 10.3389/fpsyg.2014.00830

Emmorey, K., McCullough, S., Mehta, S., and Grabowski, T. J. (2014). How sensory-motor systems impact the neural organization for language: direct contrasts between spoken and signed language. *Front. Psychol.* 5:484. doi: 10.3389/fpsyg.2014.00484

Fay, N., Lister, C. J., Ellison, T. M., and Goldin-Meadow, S. (2014). Creating a communication system from scratch: gesture beats vocalization hands down. *Front. Psychol.* 5:354. doi: 10.3389/fpsyg.2014.00354

Guellai, B., Langus, A., and Nespor, M. (2014). Prosody in the hands of the speaker. *Front. Psychol.* 5:700. doi: 10.3389/fpsyg.2014.00700

Kastner, I., Meir, I., Sandler, W., and Dachkovsky, S. (2014). The emergence of embedded structure: insights from kafr qasem sign language. *Front. Psychol.* 5:525. doi: 10.3389/fpsyg.2014.00525

Kocab, A., Pyers, J., and Senghas, A. (2015). Referential shift in nicaraguan sign language: a transition from lexical to spatial devices. *Front. Psychol.* 5:1540. doi: 10.3389/fpsyg.2014.01540

Lillo-Martin, D., de Quadros, R. M., Chen Pichler, D., and Fieldsteel, Z. (2014). Language choice in bimodal bilingual development. *Front. Psychol.* 5:1163. doi: 10.3389/fpsyg.2014.01163

Morgan, G. (2014). On language acquisition in speech and sign: development drives combinatorial structure in both modalities. *Front. Psychol.* 5:1217. doi: 10.3389/fpsyg.2014.01217

Napoli, D. J., and Sutton-Spence, R. (2014). Order of the major constituents in sign languages; Implications for all language. *Front. Psychol.* 5:376. doi: 10.3389/fpsyg.2014.00376

Pinker, S. (1999). *Words and Rules: The Ingredients of Language.* New York, NY: Basic Books.

Supalla, T., Hauser, P., and Bavelier, D. (2014). Reproducing American sign language sentences: cognitive scaffolding in working memory. *Front. Psychol.* 5:859. doi: 10.3389/fpsyg.2014.00859

Woll, B. (2014). Moving from hand to mouth: echo phonology and the origins of language. *Front. Psychol.* 5:662. doi: 10.3389/fpsyg.2014.00662

The role of syllables in sign language production

Cristina Baus[1]*, Eva Gutiérrez[2] and Manuel Carreiras[3,4,5]

[1] Laboratoire de Psychologie Cognitive, Centre National de la Recherche Scientifique (CNRS), Université d'Aix-Marseille, Marseille, France
[2] Deafness, Cognition and Language Research Centre, University College London, London, UK
[3] BCBL - Basque Research Center on Cognition, Brain and Language, Donostia, Spain
[4] IKERBASQUE, Basque Foundation for Science, Bilbao, Spain
[5] Departamento de Lengua Vasca y Comunicación, Universidad del País Vasco, Donostia, Spain

The aim of the present study was to investigate the functional role of syllables in sign language and how the different phonological combinations influence sign production. Moreover, the influence of age of acquisition was evaluated. Deaf signers (native and non-native) of Catalan Signed Language (LSC) were asked in a picture-sign interference task to sign picture names while ignoring distractor-signs with which they shared two phonological parameters (out of three of the main sign parameters: *Location*, *Movement*, and *Handshape*). The results revealed a different impact of the three phonological combinations. While no effect was observed for the phonological combination Handshape-Location, the combination Handshape-Movement slowed down signing latencies, but only in the non-native group. A facilitatory effect was observed for both groups when pictures and distractors shared Location-Movement. Importantly, linguistic models have considered this phonological combination to be a privileged unit in the composition of signs, as syllables are in spoken languages. Thus, our results support the functional role of syllable units during phonological articulation in sign language production.

Keywords: sign language, speech production, syllables, sign parameters, picture naming

INTRODUCTION

In recent years, research on sign language has accumulated evidence to suggest that spoken and sign languages are governed by similar cognitive mechanisms and underpinned by similar neuroanatomical substrates. For instance, the existence of the same linguistic phenomena in both modalities has been taken as evidence that levels of linguistic processing (semantic, lexical, and phonological) are modality-independent. The same semantic, lexical, and phonological effects reported in the spoken modality have been replicated in the sign modality (e.g., Emmorey and Corina, 1990; Corina and Knapp, 2006; Baus et al., 2008; Carreiras et al., 2008; Gutierrez et al., 2012a,b; Hosemann et al., 2013; see Carreiras, 2010 for a review). Furthermore, the same left-lateralized brain network has been described to underlie the processing of signed and spoken languages (e.g., San Jose-Robertson et al., 2004; Emmorey et al., 2007; see also MacSweeney et al., 2008 for a review).

Signs, as well as words, can be decomposed into minimal phonological constituents or formational parameters (Emmorey, 2002; but see Johnston and Schembri, 1999). Three have been considered the main formational parameters of signs (Stokoe, 1960): the *Location* of the sign in relation to the body, the *Movement* of the hand/s and the *Handshape*. Importantly, different studies suggest that these parameters play a different role during language processing (see also current phonological models in sign language; e.g., Brentari, 1998). For instance, using a picture-sign interference task, Baus et al. (2008) reported that

lexical access was facilitated when the sign corresponding to the picture and the distractor-sign shared the Handshape, while it was hampered when the Location was shared (see also, Corina and Hildebrandt, 2002; Carreiras et al., 2008; Gutierrez et al., 2012b, for similar results in sign comprehension; see Caselli and Cohen-Goldberg, 2014, for a computational model). Despite the importance of these results, sign production research is still very scarce and hence more evidence is necessary to characterize the role of these phonological parameters and the possible interactions among them. In the present study, we aimed to understand better the processes underlying sign production by asking whether phonological constituents (Location, Movement, and Handshape) are combined into higher order units before a sign is articulated, as phonemes are combined into syllables in spoken languages. To that end, the impact of the different combinations of phonological parameters on sign production was explored.

In spoken languages, syllables are considered the functional units during speech planning (e.g., Levelt and Wheeldon, 1994; Carreiras and Perea, 2004; Cholin et al., 2006; Laganaro and Alario, 2006). Accordingly, models of speech production describe the locus of syllables within the production system, either at the word-form encoding level (phonological syllables, see Dell, 1988) or during articulatory preparation (e.g., Crompton, 1981; Levelt and Wheeldon, 1994). Experimental evidence for the existence of syllables in speech production comes from different sources, such as speech errors or syllabic effects. For instance, it has been

shown that speech errors respect the *syllable position constraint* (e.g., Boomer and Laver, 1968; Mackay, 1970). That is, for those sound/form exchanges occurring between close-by words (such as rack pat for pack rat), onsets are exchanged with onsets but not with codas. Moreover, the role of syllabic units in word production has been explored mainly through two effects: the so-called *syllabic frequency effect* and the *syllabic priming effect*. The syllabic frequency effect refers to the observation that speakers are faster at naming words (and pseudowords) containing high frequency syllables than low frequency ones (e.g., Levelt and Wheeldon, 1994; Aichert and Ziegler, 2004; Alario et al., 2004; Carreiras and Perea, 2004; Cholin et al., 2006; Laganaro and Alario, 2006). The syllabic-priming effect refers to the observation that speakers are faster at naming a word (e.g., basis) when it has been primed with a syllable (ba) that respects the syllable boundaries of the word, than with an incongruent syllable (bas) that does not respect such boundaries (e.g., Ferrand et al., 1996, 1997; but see, Baumann, 1995; Schiller, 1998; Schiller et al., 2002; Schiller and Costa, 2006, for failed attempts to replicate the syllabic priming effect).

Linguistic theories of the structure of signed language agree on the existence of such syllabic-like units in signed language. That is, the syllable as a formal concept has an analog in signed language. The parallelism between syllables in spoken and signed languages stems from the idea that the way phonological constituents are organized into syllables depends on the sonority of the segments (Perlmutter, 1992). Signs are sequentially organized in terms of static-dynamic alternation that could be compared to consonants (holds) and vowels (movements) in the spoken modality. Syllables must include a *nucleus*, which corresponds to the maximal peak of sonority, the vowel, and may include an *onset* or a *coda* (Selkirk, 1982). The same applies to sign language. Models of sign language tend to attribute to the Movement the status of the nucleus (e.g., Chinchor, 1978; Brentari, 1990; Corina, 1990; Sandler, 1993; Brentari, 1998). In fact, Sandler's Location-Movement-Location model (Sandler, 1987, 1989) proposes that it is the combination of Locations and Movements that composes a syllable (see Chinchor, 1978; Wilbur, 1993, for a fairly different view). Indeed, the Movement is considered the visual equivalent of "*sonority*," being then the most salient parameter, which can be easily differentiated from the other parameters. For instance, as do vowels in the spoken modality, Movements in a sign carry prosodic as well as emotional information. Moreover, for some signs, the number of Movement repetitions determines whether a given sign is a noun or a verb (e.g., GLASS and TO DRINK in LSC have the same C Handshape next to the mouth with one repetition of the sign glass and two for to drink). Importantly however, as indicated by Emmorey et al. (2007), the fact that words and signs can be decomposed into similar syllabic-units is not a guarantee that syllabification processes are the same in word and sign production. There are several differences between spoken and signed languages that could contribute to the suggested difference in processing (for instance, most signs are monosyllabic, Brentari, 1990). Indeed, the same happens if we consider the role of syllables across different spoken languages. For instance, while syllables exist across languages, their impact as segmentation units is stronger for those languages with clear syllabic boundaries (e.g., Romance languages). Moreover, planning units

might vary depending on the task in hand. In Chinese, while syllables are the functional unit during speech production (Chen et al., 2002), logographemes are the proximal unit in handwritten production (Chen and Cherng, 2013). Thus, even if syllables have been linguistically described in signed language, it is important to describe their psychological reality by exploring how signers process these syllabic units in sign language (see Corina et al., 2014).

To date however, the functional role of syllables in sign language processing has scarcely been investigated (Corina and Knapp, 2006; Dye and Shih, 2006; Mayberry and Witcher, 2006; Gutierrez, 2008). Interestingly, these few results point to a special status of the combination of Location-Movement in both sign comprehension and production. For instance, Dye and Shih (2006) tested the speed with which deaf signers took lexical decisions in a priming paradigm in which primes and targets shared two out of the three phonological parameters (Location, Handshape, and Movement). Their results revealed that native deaf signers were faster at making decisions on the target, exclusively when prime and target shared Location and Movement. Similarly, Corina and Knapp (2006) reported a facilitatory effect for the combination Location-Movement in ASL sign production using a picture-sign interference task. However, although these results provide evidence of the privileged status of this phonological combination in sign production, they remain silent about the role of the other phonological combinations. Thus, the present study aimed to further investigate how the different combinations of parameters, namely Location-Movement, Location-Handshape and Handshape-Movement affect the speed with which signs are produced.

Our second aim was to expand Corina and Knapp's results (2006) by exploring the influence of age of acquisition on the processing of these syllabic-like units. Age of acquisition is a very interesting issue to address here, since signed language offers the unique opportunity to test age of acquisition differences in first language processing. Several studies have reported differences in performance between signers who acquire a sign language early relative to those who acquire sign language later in life (Mayberry and Fisher, 1989; Newport, 1990; Corina and Hildebrandt, 2002; Carreiras et al., 2008; Gutierrez et al., 2012a). Such differences have been attributed to a "phonological bottleneck" by which the form-based properties of signs are processed less efficiently the later the sign language is acquired. For instance, in Dye and Shih (2006), no phonological effect was observed for the Location-Movement combination when non-native signers were tested in the priming experiment. Instead, priming effects arose uniquely when primes and targets shared the Movement parameter in isolation. To further explore how age of acquisition influences lexical access during sign production, we compared the performance of two groups of signers that differed in the age at which sign language was acquired. The hypothesis is that if non-native signers are less efficient in processing phonological units in sign language, it is possible that the different phonological combinations do not equally impact native and non-native processing.

In the present study we used a picture-sign interference task (Corina and Knapp, 2006; Baus et al., 2008) and asked deaf signers who had acquired the signed language early (born within

deaf families) or late (after the age of 10) to sign the corresponding picture-sign while ignoring a distractor. The task was an adaptation of the picture-word interference paradigm, which has been extensively used in the language production literature to reveal the functional dynamics of lexical retrieval processes in speech production. Note that this is not to say that comprehension mechanisms are not involved in the processing of the distractors.
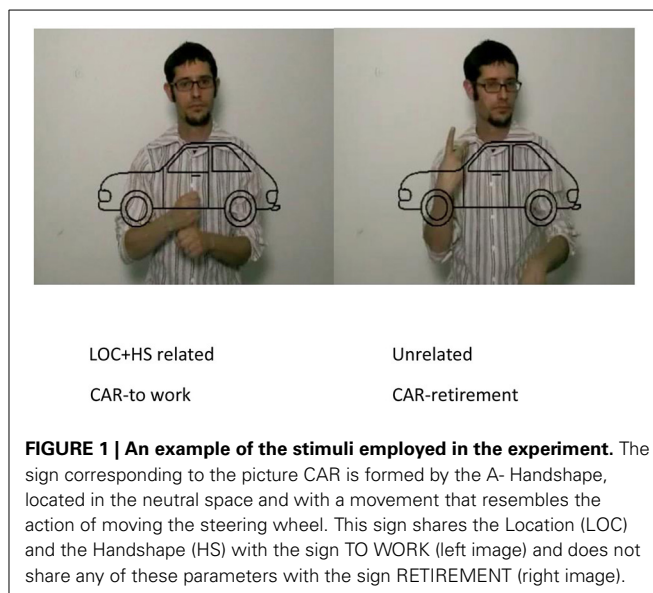
METHODS

PARTICIPANTS

Twenty-four deaf signers participated in this study (11 women). The participants were the same as in Baus et al. (2008). All of them were deaf from birth and used Catalan Sign Language (LSC) on a daily basis as preferred means of communication. Twelve participants were considered as native signers (age range 18–51, mean age 30.3, $SD = 7.6$). They were born in deaf families (parents or older siblings) and acquired the signed language before the age of 5. The remaining were non-native signers (from hearing families) (age range 18–44, mean age 26.4, $SD = 5.8$) who learned LSC at the mean age of 12 (age of exposure range 10–31 years, mean = 16, $SD = 7.2$). Both groups of participants had attended "oralist" schools (it is relatively new to find schools adapted to the deaf community). All of them had completed the years of compulsory education (primary school, up to 14 years old), with only a few of them completing the secondary levels of education (5 participants). All participants reported feeling more comfortable using the signed than the spoken language.

MATERIALS

Thirty line-drawings depicting simple objects from different semantic categories were selected (e.g., Snodgrass and Vanderwart, 1980). For each picture, two video-signs (distractors) were created: one phonologically related and one unrelated. In the phonologically related condition, the sign corresponding to the picture and the distractor-sign shared two out of the three main parameters. Thus, there were three types of phonological overlap: Handshape-Movement, Location-Handshape, and Location-Movement (ten items per condition). Given that the pool of picturable stimuli is limited, it was not possible to pair each picture-sign with a distractor-sign of each phonological condition. Thus, each picture was assigned to just one of the phonological conditions and was paired with one phonologically related and one phonologically unrelated distractor. In the unrelated condition, the picture's corresponding sign and the video-sign had no phonological or semantic relationship.

During the experiment, participants saw each picture twice, once in a phonologically related pair and once in an unrelated pair. The order of appearance was randomized. The results were then based on the comparison between the related and the unrelated conditions, where the same picture was used (see **Figure 1** for an example and the Appendix for the full list of materials in the Supplementary Material) and not on the comparison between the different phonological combinations. The pictures appeared superimposed on a video of a deaf person signing and were presented to participants at the same time (SOA 0).

FIGURE 1 | An example of the stimuli employed in the experiment. The sign corresponding to the picture CAR is formed by the A- Handshape, located in the neutral space and with a movement that resembles the action of moving the steering wheel. This sign shares the Location (LOC) and the Handshape (HS) with the sign TO WORK (left image) and does not share any of these parameters with the sign RETIREMENT (right image).

All videos had an approximate duration of 500 ms and comprised both the video distractor and the picture, that is, the picture appeared simultaneously with the onset of the distractor video sequence and remained visible on the screen together with the last frame of the video distractor until participants responded.

PROCEDURE

Participants were tested in a quiet room, avoiding visual distractors. Before the experiment started, instructions were signed to the participant in LSC. They were instructed to sign the name of the picture while ignoring the video presented at the back. After ensuring that participants understood the instructions, they were presented with a booklet containing all the pictures of the experiment to ensure that they used the designated sign during the experiment. Participants were then familiarized with the task in 10 practice trials with similar characteristics to the experimental ones.

During the experiment, the structure of the trial was as follows: (1) an instruction indicating that a new trial was about to start appeared on the screen, indicating that participants should press the two response buttons (in the response box) with their two hands and hold them pressed until their response; (2) while they pressed the response buttons, an asterisk appeared in the center of the screen for 500 ms, followed by a blank interval of 300 ms; (3) a video appeared containing the video-distractor and the picture (see **Figure 1**) and lasted for approximately 500 ms. When the video finished, the image remained still on the last frame until the participant's response; (4) 2000 ms after the participant's response, the message telling the participant to press the button responses appeared again. Reaction times were registered from the onset of the picture + video presentation to the moment the participant raised her hands off the button box to sign the name of the picture. Stimulus presentation and reaction times were controlled by Psyscope software (Cohen et al., 1993). Participants were videotaped during the experimental session to score for errors.

RESULTS

Responses different from the ones designated by the experimenter were considered as production errors and were excluded from the latency analyses. Moreover, those responses in which the participant stopped before signing were considered as hesitations and therefore counted as errors. Two pictures were also excluded because more than 80% of the participants used a sign different from the one designated by the experimenter (one picture in the Location-Handshape and one in the Location-Movement condition; indicated by an asterisk in the Appendix—Supplementary Material). Finally, signing latencies above or below two standard deviations in each condition were also excluded. Data trimming led to final exclusion of 10% of the data from the latency analysis.

Median latencies and error rates were analyzed for each phonological condition separately (Handshape-Movement, Location-Handshape, and Location-Movement). Note that using the median instead of the mean is a common practice in the analysis of populations in which a lot of variability and extreme values can be encountered.

In a 2 × 2 ANOVA, the phonological relationship (related vs. unrelated) and the group of participants (native vs. non-native) were entered as within participant and between items factors, respectively. The analyses considering the error rates did not reveal any significant results (all p's > 0.2) and they are not further discussed. Moreover, native and non-native signers did not differ in their overall signing performance. The main effect of group was not significant in any of the conditions explored (all F's < 1).

Regarding signing latencies (**Table 1**), participants were slower signing those pictures sharing Handshape and Movement with the video-distractor than the same pictures when the distractor was phonologically unrelated [$F_{1(1, 22)} = 7.47$, $p < 0.05$ and $F_{2(1, 18)} = 4.34$, $p = 0.05$]. That is, the Handshape-Movement phonological combination revealed an interference effect. Moreover, the interaction between phonological relatedness and group of participants (Natives vs. Non-natives) was significant in the analysis by participants [$F_{1(1, 22)} = 4.04$, $p = 0.05$ and $F_{2(1, 18)} = 1.29$, $p = 0.27$]. *Post-hoc* comparisons indicated that the non-native group was affected by the Handshape-Movement phonological overlap between the picture and the distractor [$F_{1(1, 22)} = 11.2$, $p < 0.01$], but not the native group ($F < 1$).

For the Location-Handshape condition (LH), there were no significant differences between the signing latencies in the related and the unrelated conditions [$F_{1(1, 22)} = 1.69$, $p = 0.20$ and $F_{2(1, 16)} = 1.72$, $p = 0.20$]. Moreover, as indicated by the lack of interaction with age of acquisition ($F < 1$), neither native nor non-natives were affected by the Location-Handshape phonological overlap.

Finally, we found a main effect of the Location-Movement combination [$F_{1(1, 22)} = 5.61$, $p < 0.05$ and $F_{2(1, 16)} = 4.41$, $p < 0.05$]. Participants were faster signing pictures when sharing the Location and the Movement with the distractor than when signing the same pictures when presented with an unrelated distractor. Both groups of participants benefited from the Location-Movement phonological overlap between target and distractor, as indicated by the lack of interaction between the phonological condition and group of participants ($F < 1$).

Table 1 | Median reaction times (RT) and percentage of errors (%error) in each phonological condition for the native and non-native group of participants.

Type of relationship		Natives			Non-natives		
		RT	**SD**	**% Error**	**RT**	**SD**	**% Error**
HM	Related	606	224	3.8	535	169	8.3
	Unrelated	596	230	4.5	474	142	5.8
	HM effect	10			59		
LH	Related	597	263	2.8	508	157	2
	Unrelated	577	215	2.5	491	152	2.7
	LH effect	20			17		
LM	Related	565	237	4.6	449	124	6.1
	Unrelated	592	256	6.5	509	180	3.0
	LM effect	−27			−50		

HM, Handshape-Movement; LH, Location-Handshape; LM, Location-Movement.

DISCUSSION

This study aimed to explore the role of the different syllabic units during sign production. Specifically, we tested whether the combination of Location and Movement, suggested by sign language models as the most important syllabic unit, would stand out during on-line LSC sign production in comparison to other parameter combinations.

Our results were clear-cut: both native and non-native signers were faster at signing the intended target only when it was presented together with a distractor that shared the Location and the Movement[1]. In line with previous research (Corina and Knapp, 2006), the present results support the idea that the combination of parameters Location-Movement seems to enjoy a privileged status during sign production, as well as during sign comprehension (e.g., Dye and Shih, 2006). Indeed, linguistic models of sign structure have described Movements and Locations as the main syllabic building blocks (e.g., Sandler, 1987; Corina and Emmorey, 1993; Brentari, 1998) with Handshapes being represented on a separate structural tier (e.g., Sandler, 1993). Although those models were created to describe signs in American Signed Language (ASL), our results and others suggest a more general effect of the Location-Movement combination across the world's signed languages, at least in what concerns Spanish Signed language (Gutierrez, 2008), British Signed Language (Dye and Shih, 2006), and Catalan Signed language. Note, however, that with these results we cannot attribute to the Location-Movement combination the unique status of syllabic unit in signed language. The reason is that finding that the Location-Movement combination influences sign production does not demonstrate that other syllabic structures do not exist in sign language (e.g., Chinchor, 1978). For instance, the Handshape-Movement combination also influenced sign production (although in the opposite direction) of non-natives, suggesting a different impact of the

---

[1]Note that although signing latencies are measured in the picture-signing interference task, both comprehension and production mechanisms are involved when performing such task.

three phonological combinations rather than the unique existence of Location-Movement as syllabic unit. Thus, the interesting question for us is: What is special about the Location-Movement combination in sign language processing? If we consider that the inventory of Locations and Movements within signed languages is significantly smaller than the inventory of Handshapes, one possibility is that particular Locations and Movements appear more frequently in the lexicon than Handshapes do. Indeed, children acquire control of the Location and Movement parameters much earlier than they master Handshapes, which require specialized dexterity of the hands and fingers (e.g., Siedlecki and Bonvillian, 1993; Conlin et al., 2000; Marentette and Mayberry, 2000). Furthermore, there is evidence that when signers make an error, the probability of involving a change in Movement or Location is relatively low (8%) compared to the probability of making an error that involves a change in the Handshape (82%; Hohenberger et al., 2002; see also Orfanidou et al., 2009). Similarly, Location and Movement are less prone to errors than Handshape in aphasic signers (Corina et al., 1992; Corina, 2000). Thus, it could be argued that our results are due to Location and Movement being more strongly represented than Handshape. However, this idea is not longer tenable if we compare the influence of these parameters when presented in isolation or jointly. Many studies have reported a facilitatory effect when Location and Movement are presented jointly, both in sign comprehension and production, and regardless of the age at which sign language was acquired. In contrast, the effect of each parameter when presented in isolation is highly variable. For instance, both inhibitory (Baus et al., 2008; Carreiras et al., 2008; see also, Caselli and Cohen-Goldberg, 2014; for a computational model on the location effects) and facilitatory effects (e.g., Dye and Shih, 2006; Orfanidou et al., 2009) have been reported when in the same task Location was manipulated in isolation. Thus, our results suggest that phonological combinations involving Location-Movement are indeed an important functional unit in lexical access and not just the additive effect of sharing two parameters (Wilbur and Allen, 1991).

Phonological combinations involving Location and Movement in sign languages have been considered to be more perceptually salient than those involving Handshape (e.g., Klima and Bellugi, 1979; Corina and Emmorey, 1993; Hohenberger et al., 2002). For instance, Hildebrant and Corina (2002) asked participants to judge the phonological similarity between a target-sign and surrounding flanker-signs, which could share the Handshape-Movement, the Location-Handshape or the Movement-Location parameters. Native signers rated those flankers that shared the Location-Movement combination more similar to the target than those involving the Handshape. Our results are in line with the idea of Location-Movement being the most salient sub-lexical (syllabic or not) unit in sign production. In this context, accessing the phonological codes composing the picture's corresponding sign would be faster for those signs sharing Location and Movement, since they will be judged as more similar than the other two phonological combinations. This would support the idea that linguistic distinctions are based on salient perceptual distinctions (Corina et al., 2014). Alternatively (but not mutually exclusive), our results could be interpreted as an effect of the *frequency* with which the parameters co-occur in sign language,

with sign-units involving Location and Movement appearing more frequently than those involving Handshape. Our results would be in line with those studies in the spoken modality showing that speakers are faster at naming words containing high-frequency syllables (which they have produced more often) than words containing low-frequency ones (e.g., Carreiras and Perea, 2004; Cholin et al., 2006; Laganaro and Alario, 2006). However, here we cannot exclude the possibility that other sublexical variables, such as the *biphone frequency* (frequency with which two phonemes co-occur regardless of whether they respect the syllabic boundaries or not), are responsible for the observed effect. In the spoken modality, the speed with which a word is produced is influenced both by the syllabic and the biphone frequency (Vitevitch et al., 2004). Such distinction has not been described in the signed modality, possibly due to the simultaneous perception of parameters within a sign. Thus, whether Location-Movement is the most frequent syllabic unit or just comprises the sequences that co-occur with more probability in the language cannot be determined from the present results. Lee and Goldrick (2008) also argued that speakers are not only sensitive to the frequency with which sub-syllabic sequences occur within a language but also to the strength of association. Importantly, if the language of the speaker determines the preference for one sequence (for instance, in Korean, sequences involving onset-vowels are strongly associated, whereas in English it is vowel-coda sequences which are more associated), it is possible that our results reveal the preference of signers for those sequences strongly associated in sign language, namely Location-Movement sequences. At present, we cannot determine whether the origin of the observed effect stems from Location-Movement being the most salient structure or the phonological sequence more probable in the language, but this opens interesting questions for future studies on phonological processing in signed language.

Finally, regarding the question of how the age of sign language acquisition might influence its phonological processing, we did not find differences between groups for the Location and Movement combination. However, the two groups differed in two aspects. Firstly, there was a tendency for shorter latencies in the non-native group than in the native one. This result was unexpected if we consider previous evidence pointing to less efficient phonological processing in non-native signers (e.g., Gutierrez et al., 2012a). Nevertheless, the fact that such differences were not significant, together with the observation that the non-native signers were overall younger than the native signers and that this is known to have an impact on processing-speed (e.g., Salthouse, 1993), prevent us from making further interpretations. Secondly and more interesting, we only obtained a difference between the two groups for the Handshape-Movement combination. Non-native signers were slower at signing pictures in the presence of the Handshape-Movement phonological distractor than in the presence of an unrelated distractor. This piece of evidence supports the idea that the late acquisition of signs results in subtle differences in sign language processing (Newport, 1990; Mayberry and Eichen, 1991; Neville et al., 1997; Corina and Hildebrandt, 2002; Newman et al., 2002; Carreiras et al., 2008; Morford et al., 2008) often involving a qualitatively different processing of Handshapes (Emmorey et al., 2003; Carreiras et al., 2008; Orfanidou et al., 2009; Best et al., 2010; Gutierrez et al.,

2012a). For instance, Hildebrant and Corina (2002) found that non-native signers judged signs as perceptually more similar when sharing the Handshape than the other parameters, while native signers based their decision on the Movement. However, Handshape cannot be the only explanation for two reasons: (1) a facilitatory effect of Handshape was reported by Baus et al. (2008) while the manipulation of Handshape in combination with Movement led to an interference effect, and (2) if Handshape is the parameter responsible for the pattern of results found for non-natives, similar results would be expected for the other phonological combination involving Handshape, namely the Location-Handshape condition. Considering these and previous findings, the pattern of results reported for non-natives is rather complex, even when more sensitive techniques such as ERPs have been employed (Gutierrez et al., 2012a). For instance, Gutierrez et al. (2012a) found that non-natives were not affected by Handshape relatedness during sign recognition either of signs or non-signs, while previous studies have reported Handshape to be the most salient phonological parameter for late signers (Corina and Hildebrandt, 2002). Thus, at this point, any interpretation of the interference effect observed for Handshape-Movement would be very tentative and premature, but it opens an excellent question to pursue in the future. Importantly, this effect also supports the idea indicated above that two-parameter effects are not just the additive effect of the two single parameters.

CONCLUSION

In sum, our results provide clear evidence of the special role that certain phonological combinations play in sign language production. Location-Movement is the only phonological combination that enjoys a benefit in processing during sign production.

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://www.frontiersin.org/journal/10.3389/fpsyg.2014.01254/abstract

REFERENCES

Aichert, I., and Ziegler, W. (2004). Syllable frequency and syllable structure in apraxia of speech. *Brain Lang.* 88, 148–159. doi: 10.1016/S0093-934X(03)00296-7

Alario, F.-X., Ferrand, L., Laganaro, M., New, B., Frauenfelder, U. H., and Segui, J. (2004). Predictors of picture naming speed. *Behav. Res. Methods Instrum. Comput.* 36, 140–155. doi: 10.3758/BF03195559

Baumann, M. (1995). *The Production of Syllables in Connected Speech.* Ph.D. dissertation, Nijmegen University.

Baus, C., Gutiérrez-Sigut, E., Quer, J., and Carreiras, M. (2008). Lexical access in Catalan Signed Language (LSC) production. *Cognition* 108, 856–865. doi: 10.1016/j.cognition.2008.05.012

Best, C., Mathur, G., Miranda, K., and Lillo-Martin, D. (2010). Effects of sign language experience on categorical perception of dynamic ASL pseudosigns. *Atten. Percept. Psychophys.* 72, 747–762. doi: 10.3758/APP.72.3.747

Boomer, D. S., and Laver, J. D. M. (1968). Slips of the tongue. *Br. J. Disord. Commun.* 3, 2–12. doi: 10.3109/13682826809011435

Brentari, D. (1990). *Theoretical Foundations of American Sign Language Phonology.* Unpublished doctoral dissertation, University of Chicago.

Brentari, D. (1998). *A Prosodic Model of Sign Language Phonology.* Cambridge, MA: MIT Press.

Carreiras, M. (2010). Sign language processing. *Lang. Linguist. Compass* 4, 430–444. doi: 10.1111/j.1749-818X.2010.00192.x

Carreiras, M., Gutiérrez-Sigut, E., Baquero, S., and Corina, D. (2008). Lexical processing in Spanish Sign Language (LSE). *J. Mem. Lang.* 58, 100–122. doi: 10.1016/j.jml.2007.05.004

Carreiras, M., and Perea, M. (2004). Naming pseudowords in Spanish: effects of syllable frequency. *Brain Lang.* 90, 393–400. doi: 10.1016/j.bandl.2003.12.003

Caselli, N. K., and Cohen-Goldberg, A. M. (2014). Lexical access in sign language: a computational model. *Front. Psychol.* 5:428. doi: 10.3389/fpsyg.2014.00428

Chen, J. Y., Chen, T. M., and Dell, G. (2002). Word-Form encoding in Mandarin Chinese as assessed by the implicit priming task. *J. Mem. Lang.* 46, 751–781. doi: 10.1006/jmla.2001.2825

Chen, J. Y., and Cherng, R. J. (2013). The proximate unit in Chinese handwritten character production. *Front. Psychol.* 4:517. doi: 10.3389/fpsyg.2013.00517

Chinchor, N. (1978). "The Syllable in ASL," in *Paper Presented at the MIT Sign Language Symposium* (Cambridge, MA).

Cholin, J., Levelt, W. J. M., and Schiller, N. O. (2006). Effects of syllable frequency in speech production. *Cognition* 99, 205–235. doi: 10.1016/j.cognition.2005.01.009

Cohen, J. D., MacWhinney, B., Flatt, M., and Provost, J. (1993). PsyScope: a new graphic interactive environment for designing psychology experiments. *Behav. Res. Methods Instrum. Comput.* 25, 257–271. doi: 10.3758/BF03204507

Conlin, K. E., Mirus, G. R., Mauk, C., and Meier, R. P. (2000). "The acquisition of first signs: place, handshape, and movement," in *Language Acquisition by Eye*, eds C. Chamberlain, J. P. Morford, and R. I. Mayberry (Mahwah, NJ: Erlbaum), 51–70.

Corina, D. (1990). Reassessing the role of sonority in syllable structure: evidence from a visual-gestural language. *CLS* 26, 33–44.

Corina, D. (2000). "Some observations regarding paraphasia in American Sign Language," in *The Signs of Language Revisited: An Anthology to Honor Ursula Bellugi and Edward Klima*, eds K. Emmorey and H. Lane (Mahwah, NJ: Erlbaum), 493–507.

Corina, D. P., and Emmorey, K. (1993). "Lexical priming in American Sign Language," in *Paper Presented at the Linguistic Society of American Conference* (Philadelphia, PA).

Corina, D. P., Hafer, S., and Welch, K. (2014). Phonological awareness for American Sign Language. *J. Deaf Stud. Deaf Educ. Adv.* 19, 530–545. doi: 10.1093/deafed/enu023

Corina, D. P., and Hildebrandt, U. (2002). "Psycholinguistic investigations of phonological structure in American Sign Language," in *Modality and Structure in Signed and Spoken Languages*, eds R. P. Meier, K. Cormier, and D. Quinto-Pozos (Cambridge: Cambridge University Press), 88–111.

Corina, D. P., and Knapp, H. P. (2006). "Lexical retrieval in American Sign Language production," in *Laboratory Phonology, 8: Varieties of Phonological Competence*, eds L. Goldstein, D. Whalen, and C. Best (Berlin: Mouton de Gruyter), 213–239.

Corina, D. P., Poizner, H., Bellugi, U., Feinberg, T., Dowd, D., and O'Grady-Batch, L. (1992). Dissociation between linguistic and nonlinguistic gestural systems: a case for compositionality. *Brain Lang.* 43, 414–447. doi: 10.1016/0093-934X(92)90110-Z

Crompton, A. (1981). Syllables and segments in speech production. *Linguistics* 19, 663–716. doi: 10.1515/ling.1981.19.7-8.663

Dell, G. S. (1988). The retrieval of phonological forms in production: tests of predictions from a connectionist model. *J. Mem. Lang.* 27, 124–142. doi: 10.1016/0749-596X(88)90070-8

Dye, M. W. G., and Shih, S. (2006). "Phonological priming in British Sign Language," in *Laboratory Phonology, 8: Varieties of Phonological Competence*, eds L. Goldstein, D. Whalen, and C. Best (Berlin: Mouton de Gruyter).

Emmorey, K. (2002). *Language, Cognition, and the Brain: Insights from Sign Language Research*. Mahwah, NJ: Erlbaum.

Emmorey, K., and Corina, D. (1990). Lexical recognition in sign language: effects of phonetic structure and morphology. *Percept. Mot. Skills* 71, 1227–1252.

Emmorey, K., McCullough, S., and Brentari, D. (2003). Categorical perception in American Sign Language. *Lang. Cogn. Process.* 18, 21–45. doi: 10.1080/01690960143000416

Emmorey, K., Mehta, S., and Grabowski, T. J. (2007). The neural correlates of sign and word production. *Neuroimage* 36, 202–208. doi: 10.1016/j.neuroimage.2007.02.040

Ferrand, L., Segui, J., and Grainger, J. (1996). Masked priming of word and picture naming: the role of syllable units. *J. Mem. Lang.* 35, 708–723. doi: 10.1006/jmla.1996.0037

Ferrand, L., Segui, J., and Humphreys, G. W. (1997). The syllable's role in word naming. *Mem. Cognit.* 25, 458–470. doi: 10.3758/BF03201122

Gutierrez, E. (2008). *El Papel de Los Parámetros Fonológicos en el Procesamiento de la Lengua de Signos Española*. Unpublished doctoral dissertation, University of la Laguna, Tenerife, Spain.

Gutierrez, E., Müller, O., Baus, C., and Carreiras, M. (2012a). Electrophysiological evidence for phonological priming in Spanish Sign Language lexical access. *Neuropsychologia* 50, 1335–1346. doi: 10.1016/j.neuropsychologia.2012.02.018

Gutierrez, E., Williams, D., Grosvald, M., and Corina, D. (2012b). Lexical access in American Sign Language: an ERP investigation of effects of semantics and phonology. *Brain Res.* 1468, 63–83. doi: 10.1016/j.brainres.2012.04.029

Hildebrant, U., and Corina, D. P.(2002). Phonological similarity in American Sign Language. *Lang. Cogn. Process.* 6, 593–612. doi: 10.1080/01690960143000371

Hohenberger, A., Happ, D., and Leuninger, H. (2002). "Modality dependent aspects of signed language production: evidence from slips of the hands and their repairs in German Sign Language," in *Modality and Structure in Signed and Spoken Language*, eds R. P. Meier, K. Cormier, and D. Quinto-Pozos (Cambridge: Cambridge University Press), 112–142.

Hosemann, J., Hermann, A., Steinbach, M., Bornkessel-Schlesewsky, I., and Schlesewsky, M. (2013). Lexical prediction via forward models: N400 evidence from German Sign Language. *Neuropsychologia* 51, 2224–2237. doi: 10.1016/j.neuropsychologia.2013.07.013

Johnston, T., and Schembri, A. (1999). On defining lexeme in a Signed Language. *Sign Lang. Linguist.* 2, 115–185. doi: 10.1075/sll.2.2.03joh

Klima, E. S., and Bellugi, U. (1979). *The Signs of Language*. Cambridge, MA: Harvard University Press.

Laganaro, M., and Alario, F.-X. (2006). On the locus of the syllable frequency effect in speech production. *J. Mem. Lang.* 55, 178–196. doi: 10.1016/j.jml.2006.05.001

Lee, Y., and Goldrick, M. (2008). The emergence of sub-syllabic representations. *J. Mem. Lang.* 59, 155–168. doi: 10.1016/j.jml.2008.03.002

Levelt, W. J. M., and Wheeldon, L. (1994). Do speakers have access to a mental syllabary? *Cognition* 50, 239–269.

Mackay, D. G. (1970). Spoonerisms: the structure of errors in the serial order of speech. *Neuropsychologia* 8, 323–350. doi: 10.1016/0028-3932(70)90078-3

MacSweeney, M, Capek, C. M., Campbell, R., and Woll, B. (2008). The signing brain: the neurobiology of sign language. *Trends Cogn. Sci.* 12, 432–440. doi: 10.1016/j.tics.2008.07.010

Marentette, P. F., and Mayberry, R. (2000). "Principles for an emerging phonological system: a case study of acquisition of early ASL," in *Language Acquisition by Eye*, eds C. Chamberlain and J. P. Morford (Nawhah, NJ: Lawrence Erlbaum Associates), 71–90.

Mayberry, R., and Eichen, E. (1991). The long-lasting advantages of learning sign language in childhood: another look at the critical period for language acquisition. *J. Mem. Lang.* 30, 486–512. doi: 10.1016/0749-596X(91)90018-F

Mayberry, R. I., and Witcher, P. (2006). *What Age of Acquisition Effects Reveal About the Nature of Phonological Processing*. CRL Technical Reports, Vol. 17, University of California, San Diego, CA.

Mayberry, R. I., and Fisher, S. D. (1989). Looking through phonological shape to lexical meaning: the bottleneck of non-native sign language processing. *Mem. Cognit.* 17, 740–754.

Morford, J. P., Grieve Smith, A. B., MacFarlane, J., Staley, J., and Waters, G. S. (2008). Effects of language experience on the perception of American Sign Language. *Cognition* 109, 41–53. doi: 10.1016/j.cognition.2008.07.016

Neville, H., Coffey, S., Lawson, D., Fischer, A., Emmorey, K., and Bellugi, U. (1997). Neural systems mediating American Sign Language: effects of sensory experience and age of acquisition. *Brain Lang.* 57, 285–308. doi: 10.1006/brln.1997.1739

Newman, A., Bavelier, D., Corina, D., Jezzard, P., and Neville, H. (2002). A critical period for right hemisphere recruitment in American Sign Language processing. *Nat. Neurosci.* 5, 76–80. doi: 10.1038/nn775

Newport, E. (1990). Maturational constrains on language learning. *Cogn. Sci.* 14, 11–28. doi: 10.1207/s15516709cog1401_2

Orfanidou, E., Adam, R., McQueen, J. M., and Morgan, G. (2009). Making sense of nonsense in British Sign Language (BSL): the contribution of different phonological parameters to sign recognition. *Mem. Cognit.* 37, 302–315. doi: 10.3758/MC.37.3.302

Perlmutter, D. (1992). Sonority and syllable structure in American Sign Language. *Linguist. Inq.* 23, 407–442.

Salthouse, T. A. (1993). Speed and knowledge as determinants of adult age differences in verbal tasks. *J. Gerontol.* 4, 29–36. doi: 10.1093/geronj/48.1.P29

Sandler, W. (1987). Assimilation and feature hierarchy in American Sign Language. *Chic. Linguist. Soc.* 23, 266–278.

Sandler, W. (1989). *Phonological Representation of the Sign: Linearity and Nonlinearity in American Sign Language*. Dordrecht: Foris. doi: 10.1515/9783110250473

Sandler, W. (1993). A sonority cycle in American Sign Language. *Phonology* 10, 243–280. doi: 10.1017/S0952675700000051

San Jose-Robertson, L., Corina, D. P., Ackerman, D., Guillemin, A., and Braun, A. R (2004). Neural systems for sign language production: mechanisms supporting lexical selection, phonological encoding, and articulation. *Hum. Brain Mapp.* 23, 156–167. doi: 10.1002/hbm.20054

Schiller, N. O. (1998). The effect of visually masked primes on the naming latencies of words and pictures. *J. Mem. Lang.* 39, 484–507. doi: 10.1006/jmla.1998.2577

Schiller, N. O., and Costa, A. (2006). Activation of segments, not syllables, during phonological encoding in speech production. *Mental Lexicon* 1, 231–250. doi: 10.1075/ml.1.2.04sch

Schiller, N. O., Costa, A., and Colomé, A. (2002). "Phonological encoding of single words: in search of the lost syllable," in *Papers in Laboratory Phonology*, eds C. Gussenhoven and N. Warner (Berlin: Mouton de Gruyter), 35–59.

Selkirk, E. O. (1982). *The Syntax of Words*. Cambridge, MA: MIT Press.

Siedlecki, T., and Bonvillian, J. D. (1993). Location, handshape, and movement: young children's acquisition of the formational aspects of American Sign Language. *Sign Lang. Stud.* 78, 31–52. doi: 10.1353/sls.1993.0016

Snodgrass, J. G., and Vanderwart, M. (1980). A standarized set of 260 pictures: norms for name agreement, image agreement, familiarity and visual complexity. *J. Exp. Psychol. Hum. Learn. Mem.* 6, 174–215.

Stokoe, W. (1960). *Sign Language Structure: an Outline of the Visual Communication Systems of the American Deaf*. Silver Spring, MD: Linstok Press.

Vitevitch, M. S., Armbrüster, J., and Chu, S. (2004). Sublexical and lexical representations in speech production: effects of phonotactic probability and onset density. *J. Exp. Psychol. Learn. Mem. Cogn.* 30, 514–529. doi: 10.1037/0278-7393.30.2.514

Wilbur, R. (1993). "Syllables and segments: hold the movement and move the holds," in *Phonetics and Phonology, Vol 3: Current Issues in ASL Phonology*, ed G. Coulter (San Diego, CA: Academic Press), 135–166.

Wilbur, R., and Allen, G. D. (1991). Perceptual evidence against internal structure in American Signed Language. *Lang. Speech* 34, 27–46.

Prosody in the hands of the speaker

Bahia Guellaï¹*, Alan Langus² and Marina Nespor²

¹ Laboratoire Ethologie, Cognition, Développement, Département de Psychologie, Université Paris Ouest Nanterre La Défense, Nanterre, France
² Language Cognition and Development Laboratory, Cognitive Neuroscience Sector, International School for Advanced Studies, Trieste, Italy

In everyday life, speech is accompanied by gestures. In the present study, two experiments tested the possibility that spontaneous gestures accompanying speech carry prosodic information. Experiment 1 showed that gestures provide prosodic information, as adults are able to perceive the congruency between low-pass filtered—thus unintelligible—speech and the gestures of the speaker. Experiment 2 shows that in the case of ambiguous sentences (i.e., sentences with two alternative meanings depending on their prosody) mismatched prosody and gestures lead participants to choose more often the meaning signaled by gestures. Our results demonstrate that the prosody that characterizes speech is not a modality specific phenomenon: it is also perceived in the spontaneous gestures that accompany speech. We draw the conclusion that spontaneous gestures and speech form a single communication system where the suprasegmental aspects of spoken language are mapped to the motor-programs responsible for the production of both speech sounds and hand gestures.

Keywords: gestures, comprehension, speech perception, ambiguity, prosody

INTRODUCTION

Human language is a multimodal experience: it is perceived through both ears and eyes. When perceiving speech, adults automatically integrate auditory and visual information (McGurk and MacDonald, 1976), and seeing someone speaking may improve speech intelligibility (Sumby and Pollack, 1954). The visual information involved in speech is not limited to the lips, the mouth and the head, but can also involve other cues such as eyebrow movements (Bernstein et al., 1998; Graf et al., 2002; Krahmer and Swerts, 2004; Munhall et al., 2004). In fact, in face-to-face interactions people use more than their voice to communicate: the whole body is involved and may serve informative purposes (Kendon, 1994; Kelly and Barr, 1999 for a review). For example, when interacting with others, people all around the world usually also produce spontaneous gestures while talking. In fact gestures are so connected with speech that people may be found gesturing when nobody sees them (Corballis, 2002) and even congenitally blind people gesture when interacting with each other (Iverson and Goldin-Meadow, 1998). Yet, the role of gestures that accompany speech (i.e., co-speech gestures) in communication is still not well understood and little if any attention to the relation between co-speech gestures and the syntactic and prosodic structure of spoken language has been paid in previous studies. Some authors claim that these co-speech gestures are not produced to serve any communicative purposes (Rimé and Shiaratura, 1991). On the contrary, others suggest that gestures and speech are parts of the same system and are performed for the purpose of expression (Kendon, 1983; McNeill, 1992). One way to understand the implication of co-speech gestures in communication is to study their implications at the different levels of the utterance. The present study aimed to investigate the role of

gestures that accompany speech at the prosodic level in speech perception.

Gestures accompanying speech are known to ease the speaker's cognitive load, and gesturing helps solving diverse individual tasks ranging from mathematics to spatial reasoning (Cook and Goldin-Meadow, 2006; Chu and Kita, 2011). Gestures are also believed to promote learning in adults as well as in children (Ping and Goldin-Meadow, 2010), to aid the conceptual planning of messages (Alibali et al., 2000), and to facilitate lexical access (Alibali et al., 2000). This suggests that gestures that accompany speech might maximize information about events by providing it cross-modally (de Ruiter et al., 2012). In fact, human infants' canonical babbling is temporally related to rhythmic hand activity already at 30 weeks of age (Locke et al., 1995), suggesting that gestures and speech go "hand-in-hand" from the earliest stages of cognitive development (McNeill, 1992; So et al., 2009).

Here we investigate whether gestures also convey some information about the prosodic structure of spoken language. We test whether prosody, an essential aspect of language, is also detected in gestures. In the auditory modality, prosody is characterized by changes in duration, intensity and pitch (for an overview see Cutler et al., 1997; Warren, 1999; Speer and Blodgett, 2006; Langus et al., 2012). Speakers can intentionally manipulate these acoustic cues to convey information about their states of mind (e.g., irony or sarcasm), to define the type of speech act they are making (e.g., a question or an assertion), and to highlight certain elements over others (e.g., by contrasting them). Importantly, prosody also conveys information about the structure of language. Because the grammatical structure of human language is automatically mapped onto prosodic structure during speech production (Langus et al., 2012), the prosody of spoken language

also signals the grammatical structure (Nespor and Vogel[1], 1986, 2007). Though prosody offers cues to different aspects of grammar, here we concentrate on the role of prosody in conveying information about syntactic structure.

It has been observed that prosodic cues are the most reliable cues for segmenting continuous speech cross-linguistically (Cutler et al., 1997). Adult listeners can use these cues to constrain lexical access (Christophe et al., 2004), to locate major syntactic boundaries in speech (Speer et al., 2011), and to determine how these relate to each other in sentences (Fernald and McRoberts, 1995; Langus et al., 2012). This is best seen in cases where listeners can disambiguate sentences that have more than one meaning (e.g., [bad] [boys and girls] vs. [bad boys] [and girls]) by relying on prosody alone (Lehiste et al., 1976; Nespor and Vogel, 1986, 2007; Price et al., 1991). Manipulations of the prosodic structure influence how listeners interpret syntactically ambiguous utterances (Lehiste, 1973; Lehiste et al., 1976; Cooper and Paccia-Cooper, 1980; Beach, 1991; Price et al., 1991; Carlson et al., 2001; see Cutler et al., 1997). These effects of prosody emerge quickly during online sentence comprehension, suggesting that they involve a robust property of the human parser (Marslen-Wilson et al., 1992; Warren et al., 1995; Nagel et al., 1996; Pynte and Prieur, 1996; Kjelgaard and Speer, 1999; Snedeker and Trueswell, 2003; Weber et al., 2006). Naive speakers systematically vary their prosody depending on the syntactic structure of sentences and naive listeners can use this variation to disambiguate utterances that—though containing the same sequence of words—differ in that they are mapped from sentences with different syntactic structures (Nespor and Vogel, 1986, 2007; Snedeker and Trueswell, 2003; Kraljic and Brennan, 2005; Schafer et al., 2005). These studies indicate that users of spoken language share implicit knowledge about the relationship between prosody and syntax and that they can use both during speech production and comprehension. To account for the syntax-prosody mapping, Nespor and Vogel (1986, 2007) have proposed a hierarchy that at the phrasal level contains—among other constituents—the Phonological Phrase (PP) and the Intonational Phrase (IP). These constituents are signaled in different ways: besides being signaled through external sandhi rules that are bound to a specific constituent, the PP right edge is signaled through final lengthening, and the IP level is signaled through pitch resetting at the left edge and through final lengthening at the right edge.

Here we ask whether prosody could also be perceived visually in the spontaneous gestures that accompany speech. In English and Italian, specific hand gestures ending with an abrupt stop, called "beats" (i.e., McNeill, 1992), are temporally related to pitch accents in speech production (Yasinnik et al., 2004; Esposito et al., 2007; Krahmer and Swerts, 2007). Also in sign languages, prosodic cues are not only conveyed through facial expressions, but also through hand and body movements (Nespor and Sandler, 1999; Wilbur, 1999; Sandler, 2011; Dachkovsky et al., 2013). A model developed on the basis of Israeli Signed Language

showed that body positions align with rhythmic manual features of the signing stream to mark prosodic constituents' boundaries at different levels of the prosodic hierarchy (Nespor and Sandler, 1999; Sandler, 1999, 2005, 2011). More recently, Sandler (2012) proposed that many actions of the body in sign languages—that she calls "dedicated gestures"—perform linguistic functions and contribute to prosodic structure.

Do people perceive prosody and co-speech gestures as a coherent unit in everyday interactions? There is some evidence that both adults and infants match the global head and facial movements of the speaker with speech sounds (Graf et al., 2002; Munhall et al., 2004; Blossom and Morgan, 2006; Guellaï et al., 2011). However, it is unknown whether visual prosodic cues that accompany speech, but are not directly triggered by the movements of the vocal tract, are actually used to process the structure of the speech signal. Here we ask whether prosody can be perceived in the spontaneous gestures of a speaker (Experiment 1), and if listeners can use gestures to disambiguate sentences with the same sequence of words mapped onto different speech utterances that have two alternative meanings (Experiment 2). To investigate which prosodic cues participants rely on in disambiguating these sentences, we constructed sentences where disambiguation could be either due to IP or to PP boundaries. This enabled us to test whether the prosodic hierarchy is discernable from gestures alone.

EXPERIMENT 1

In this first experiment, we explored whether gestures carry prosodic information. We tested Italian-speaking participants in their ability to discriminate audio-visual presentations of low-pass filtered Italian utterances where the gestures either matched or mismatched the auditory stimuli (Singer and Goldin-Meadow, 2005). While low-pass filtering renders speech unintelligible, it preserves the prosody of the acoustic signal (Knoll et al., 2009). This guaranteed that only prosodic information was available to the listeners.

METHODS

Participants

We recruited 20 native speakers of Italian (15 females and 5 males, mean age 24 ± 5) from the subject pool of SISSA—International School of Advanced Studies (Trieste, Italy). Participants reported no auditory, vision, or language related problems. They received monetary compensation.

Stimuli

We used sentences that contain the same sequence of words and that can be disambiguated using prosodic cues at one of two different levels of the prosodic hierarchy. The disambiguation could take place at the IP level—the higher of these two constituents, coextensive with intonational contours—signaled through pitch resetting and final lengthening (Nespor and Vogel, 1986, 2007). For example, in Italian, *Quando Giacomo chiama suo fratello è sempre felice* is ambiguous because depending on the IP boundary *è sempre felice ((he) is always happy)* could refer to either *Giacomo* or *suo fratello (his brother)*: (1) [Quando Giacomo chiama]$_{IP}$ [suo fratello è sempre felice]$_{IP}$ (*When Giacomo calls him his brother is*

---

[1]Though recursive prosodic phrasal constituents have been proposed at the level of the Intonational Phrase (Ladd, 1986) we rely on the more standardly accepted view that phrasal prosody has no recursive constituents (Selkirk, 1984; Nespor and Vogel, 1986, 2007).

always happy); or (2) [Quando Giacomo chiama suo fratello]$_{IP}$ [è sempre felice]$_{IP}$ (*When Giacomo calls his brother he is always happy*).

Alternatively, the disambiguation could take place at the PP level where phrase boundaries are signaled through final lengthening. The PP extends from the left edge of a phrase to the right edge of its head in head-complement languages (e.g., Italian and English); and from the left edge of a head to the right edge of its phrase in complement-head languages (e.g., Japanese and Turkish) (Nespor and Vogel, 1986, 2007). An example of a phrase with two possible meanings is *mappe di città vecchie* that is ambiguous in Italian because depending on the location of the PP boundaries, the adjective *vecchie* (*old*) could refer to either *città* (*towns*) or *mappe* (*maps*): (1) [mappe di città]$_{PP}$ [vecchie]$_{PP}$ (*old maps of towns*); or (2) [mappe]$_{PP}$ [di città vecchie]$_{PP}$ (*maps of old towns*) (for more details see the list of the sentences ambiguous at the IP and PP levels used in Experiments 1 and 2 in **Table 1**). The presentation of the two types of sentences—those ambiguous at the IP level and those ambiguous at the PP level—was randomized across subjects.

We video recorded two native speakers of Italian—a male and a female—uttering ten different ambiguous Italian sentences (see **Table 1**). The speakers were unaware of the purpose or the specifics of the experiments. The speakers were asked to convey to an Italian listener the different meanings of the sentences using spontaneous gestures in the most natural way possible. They were video recorded under experimental conditions (i.e., not in natural setting) uttering the different sentences presented in **Table 1** with each of their two different meanings. The co-speech gestures produced contained both iconic gestures (i.e., gestures expressing some aspects of the lexical content) and beats ones (i.e., gestures linked to some prosodic aspects of the utterance) gestures (see Kendon, 1994 for a review; McNeill, 1992). The videos of the speakers were framed so that only the top of their body, from their shoulders to their waist, was visible (see **Movies S1, S2**). Thus, the mouth—i.e., the verbal articulation of the sentences—was not visible. Two categories of videos were created from these recordings using Sony Vegas 9.0 software. One category corresponded to the "matched videos" in which the speakers' gestures and their speech matched and the second category corresponded

---

Table 1 | Sentences ambiguous at the IP or PP level used in both Experiments with their prosodic parsing and their two possible meanings translated in English.

Sentences ambiguous at the Intonational Phrase level (IP)	Sentences ambiguous at the Phonological Phrase level (PP)

[[Alla conferenza]$_{PP}$ [Luciano]$_{PP}$ [ha parlato naturalmente]$_{PP}$]$_{IP}$
   At the conference Luciano has talked in a natural way
 [[Alla conferenza]$_{PP}$ [Luciano]$_{PP}$ [ha parlato]$_{PP}$]$_{IP}$[[naturalmente]$_{PP}$]$_{IP}$
   Of course Luciano talked at the conference
	[[Come hai visto]$_{PP}$]$_{IP}$ [[la vecchia]$_{PP}$ [legge]$_{PP}$ [la regola]$_{PP}$]$_{IP}$
	   As you see the old woman reads the rule
	 [[Come hai visto]$_{PP}$]$_{IP}$ [[la vecchia legge]$_{PP}$ [la regola]$_{PP}$]$_{IP}$
	   As you see the old law rules it

[[Come ti avevo detto]$_{PP}$]$_{IP}$ [quando Giorgio]$_{PP}$ [chiama]$_{PP}$]$_{IP}$ [[suo fratello]$_{PP}$ [è sempre nervoso]$_{PP}$]$_{IP}$
   As I had told you when Giorgio calls his brother he is always happy
 [[Come ti avevo detto]$_{PP}$]$_{IP}$ [quando Giorgio]$_{PP}$ [chiama]$_{PP}$][suo fratello]$_{PP}$ [è sempre nervoso]$_{PP}$]$_{IP}$
   As I had told you when Giorgio calls his brother is always happy
	[[Come sicuramente hai visto]$_{PP}$]$_{IP}$ [la vecchia]$_{PP}$ [sbarra]$_{PP}$ [la porta]$_{PP}$]$_{IP}$
	   As you for sure have seen the old lady blocks the door
	 [[Come sicuramente hai visto]$_{PP}$]$_{IP}$ [[la vecchia]$_{PP}$ [sbarra]$_{PP}$ [la porta]$_{PP}$]$_{IP}$
	   As you for sure have seen the old bar carries it

[[Come hai visto]$_{PP}$]$_{IP}$ [[quando Luca]$_{PP}$ [chiama]$_{PP}$ [il suo gatto]$_{PP}$]$_{IP}$ [è sempre felice]$_{PP}$]
   As you have seen when Luca calls his cat he is always happy
 [[Come hai visto]$_{PP}$]$_{IP}$ [[quando Luca]$_{PP}$ [chiama]$_{PP}$]$_{IP}$ [[il suo gatto]$_{PP}$ [è sempre felice]$_{PP}$]$_{IP}$
   As you have seen when Luca calls his cat is always happy
	[[Come ti avevo detto]$_{PP}$]$_{IP}$ [[quando Luca]$_{PP}$ [legge Dante]$_{PP}$ [è felice]$_{PP}$]$_{IP}$
	   As I had told you when Luca reads Dante he is happy
	 [[Come ti avevo detto]$_{PP}$]$_{IP}$ [[quando Luca]$_{PP}$ [legge]$_{PP}$ [Dante]$_{PP}$ [è felice]$_{PP}$]$_{IP}$
	   As I had told you when Luca reads Dante is happy

[[Come ti avevo detto]$_{PP}$]$_{IP}$ [[se Giacomo]$_{PP}$ [scrive bene]$_{PP}$ [è felice]$_{PP}$]$_{IP}$
   As I had told you if Giacomo writes well he is happy
 [[Come ti avevo detto]$_{PP}$]$_{IP}$ [[se Giacomo]$_{PP}$ [scrive]$_{PP}$]$_{IP}$ [[Bene]$_{PP}$ [è felice]$_{PP}$]$_{IP}$
   As I had told you if Giacomo calls Bene is happy
	[[Sanno]$_{PP}$ [tutti]$_{PP}$ [che canta solo]$_{PP}$ [se è felice]$_{PP}$]$_{IP}$
	   Everybody knows that he sings alone if he is happy
	 [[Sanno]$_{PP}$ [tutti]$_{PP}$ [che canta]$_{PP}$ [solo]$_{PP}$ [se è felice]$_{PP}$]$_{IP}$
	   Everybody knows that he sings only if he is happy

[[Sai]$_{PP}$ [che parla]$_{PP}$ [molte lingue]$_{PP}$ [naturalmente]$_{PP}$]$_{IP}$
   You know that he speaks many languages in a natural way
 [[Sai]$_{PP}$ [che parla]$_{PP}$ [molte lingue]$_{PP}$]$_{IP}$ [[naturalmente]$_{PP}$]$_{IP}$
   You of course know that he speaks many languages	

[Come ti avevo detto]$_{PP}$]$_{IP}$ [salta]$_{PP}$ [il muro]$_{PP}$ [più alto]$_{PP}$[naturalmente]$_{PP}$]$_{IP}$
   As I had told you s/he jumps over the tallest wall in a natural way
 [Come ti avevo detto]$_{PP}$]$_{IP}$ [salta]$_{PP}$ [il muro]$_{PP}$ [più alto]$_{PP}$]$_{IP}$[[naturalmente]$_{PP}$]$_{IP}$
   As I had told you of course s/he jumps over the tallest wall	

to the "mismatched videos" in which the gestures were associated with the speech sound of the same sequence of words, but with the alternative meaning. To do so, we edited the original recordings and switched the acoustic and visual stimuli. This manipulation was not perceived by the participants as reported in the debriefing session. Then the gestures signaled the opposite meaning of that is signaled by the sentence for this condition. A total of 80 videos were created (each of the sentences was uttered twice). We ensured that, in the mismatched audio-visual presentations, the left and the right edges of the gesture sequences were aligned with the left and the right edges of the utterances (see **Figure 1**). This is an important point as in sign languages manual alignment with the signing stream is quite strict (Nespor and Sandler, 1999; Sandler, 2012) and co-speech gestures in general are tightly temporally linked to speech (McNeill et al., 2000). To remove the intelligibility of speech but to preserve prosodic information, the speech sounds were low-pass filtered using Praat software with the Haan band filter (0–400 Hz). As a result it was not possible to detect from speech which of the two meanings of a sentence was intended, as reported by the participants at the end of the experiment. The resulting stimuli had the same loudness of 70 dB.

PROCEDURE

Participants were tested in a soundproof room and the stimuli were presented through headphones. They were instructed to watch the videos and answer—by pressing a key on a keyboard—whether what they saw matched or mismatched what they heard (i.e., [S] = yes or [N] = no). A final debriefing (i.e., we explained the goals of the study) ensured that none of the participants understood the meaning of the sentences.

RESULTS AND DISCUSSION

The results show that participants correctly identified the videos in which hand gestures and speech matched [$M = 81.9$, $SD = 11.03$: t-test against chance with equal variance not assumed $t_{(19)} = 12.93$, $p < 0.0001$] and those in which they did not match [$M = 69.3$, $SD = 10.17$; $t_{(19)} = 8.41$, $p < 0.0001$]. A repeated measure ANOVA with condition (Match, Mismatch) and type of prosodic contour (IP and PP) was performed on the mean percentage. The ANOVA only revealed a significant main effect for condition [$F_{(1, 19)} = 12.81$, $p = 0.002$, $\acute{\eta}^2 = 0.4$], but neither for type of prosodic contour [$F_{(1, 19)} = 1.20$, $p = 0.287$, $\acute{\eta}^2 = 0.06$] nor for an interaction of type and condition [$F_{(1, 19)} = 3.52$, $p = 0.076$, $\acute{\eta}^2 = 0.16$]. Participants answered correctly more often in the matching condition, and there are more errors for the mismatching one. In other words, they are more likely to incorrectly accept a mismatching video than to reject a matching one. A possible interpretation for this asymmetric results is that participants may detect some incoherences in the mismatching videos and these could lead them to a certain degree of uncertainty in their answers. To sum up, the results show that adult listeners detect the congruency between hand gestures and the acoustic speech signal even when only the prosodic cues are preserved in the acoustic signal (see **Figure 2**). The spontaneous gestures that accompany speech must therefore be aligned with the speech signal, suggesting a tight link between the motor-programs responsible for producing both speech and the spontaneous gestures that accompany it. The results of Experiment 1 thus show that adult listeners are sensitive to the temporal alignment of speech and the gestures that speakers spontaneously produce when they speak. In the next Experiment we asked whether the gestures that accompany speech

FIGURE 2 | Mean percentage of right answers in the match and mismatch conditions of Experiment 1. Participants' mean percentage of right answers is significantly higher in the matching condition than in the mismatching one (**$p < 0.0001$). Errors bars represent the standard deviation.

FIGURE 1 | Examples of the stimuli used in both Experiments (i.e., with speech being filtered for Experiment 1). Here the sentence is "Come hai visto quando Luca chiama il suo gatto è sempre felice." Two meanings are possible: "As you have seen when Luca calls his cat is always happy" (meaning 1) vs. "As you have seen when Luca calls his cat he is always

happy" (meaning 2). On the left, this is the matched version (i.e., the audio and the visual inputs match) whereas on the right this is the mismatched version (i.e., the audio of meaning 1 is aligned with the visual input of meaning 2). The left and right edges of gesture sequences and those of utterances were aligned.
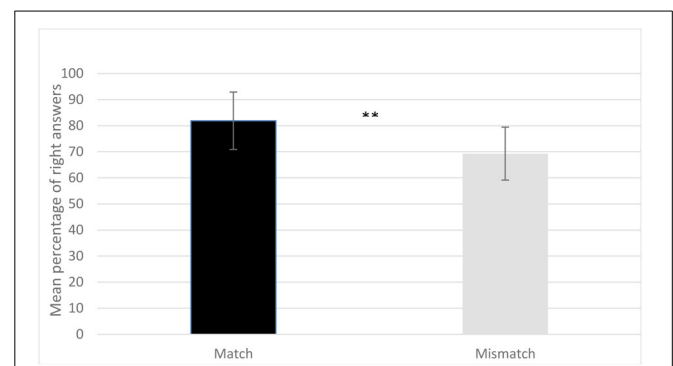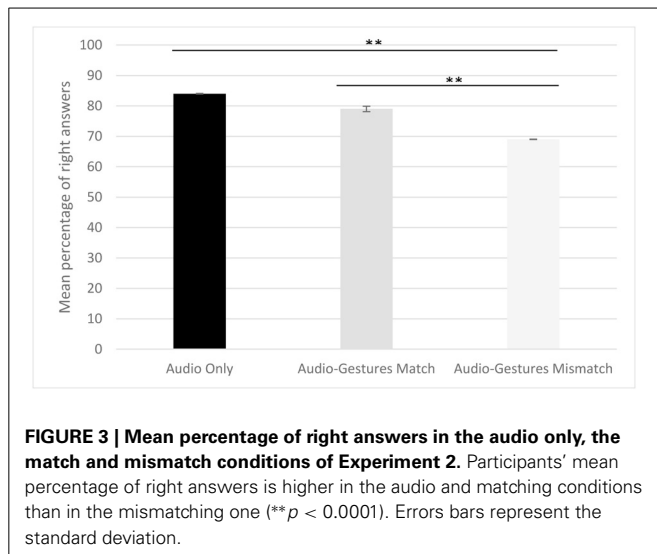
FIGURE 3 | Mean percentage of right answers in the audio only, the match and mismatch conditions of Experiment 2. Participants' mean percentage of right answers is higher in the audio and matching conditions than in the mismatching one (** $p < 0.0001$). Errors bars represent the standard deviation.

have any effect on adult listeners' understanding of ambiguous sentences.

EXPERIMENT 2

In sign languages, a good deal of prosodic information is conveyed by gestures of different parts of the face and body (Sandler, 2012). This information alone can distinguish coordinate from subordinate sentences and declarative sentences from questions (Pfau and Quer, 2010; Dachkovsky et al., 2013). This may suggest that in spoken languages too, listeners can actively use gestures accompanying speech for perceiving, processing and also understanding speech. For example, if gestures are carrying prosodic information about the grammatical structure of the speech signal, it should be easier for listeners to disambiguate a sentence that can have two different meanings when the gestures accompanying speech are visible and match the audible utterance. Experiment 2 was designed to test this hypothesis. We presented to Italian-speaking adults potentially ambiguous Italian sentences in which the audio-visual information was either matched or mismatched.

METHODS
Participants

We recruited 20 native speakers of Italian (9 females and 11 males, mean age 23 ± 3) from the subject pool of SISSA—International School of Advanced Studies (Trieste, Italy). Participants reported no auditory, vision, or language related problems. They received monetary compensation.

Stimuli

The same videos of the speakers recorded for Experiment 1 were used. However, for Experiment 2, the speech sound was not low-pass filtered (see **Movies S3**, **S4**). We added also audio-only samples of the sentences as a control condition. Thus, there were three categories of stimuli for Experiment 2: auditory only, auditory with matched gestures and auditory with mismatched gestures. For each of the categories, there were 10 different sentences (i.e., the same sentences as in Experiment 1) that could

have two different meanings, uttered by a male and a female speaker. Thus, a total of 120 stimuli were created. We ensured that the left and right edges of gesture sequences and those of utterances were aligned. Speech sounds for all the stimuli had the same loudness of 70 dB.

PROCEDURE

Participants were tested in a soundproof room with headphones. They were instructed to both listen to and to watch the stimuli. After each presentation, a question appeared on the screen regarding the meaning of the sentence they had just perceived. For example, after "Quando Giacomo chiama suo fratello è sempre felice" (When—Giacomo—calls—his—brother—is—always - happy) either the question "Giacomo è felice?" *(Is Giacomo happy?)*, or the question "Suo fratello è felice?" *(Is his brother happy?)* appeared. Participants had to answer, by clicking on a keyboard, if the answer to the question was *yes* or *no*. In each of the three within-subject conditions (audio only, audio and gestures match, audio and gestures mismatch) participants saw 5 of the 10 sentences (total 10 different meanings) so that each meaning was paired with a "yes" question ("yes" = hit/"no" = miss) and a "no" question ("yes" = correct rejection/"no" = false alarm). Each participant heard the same sentence produced by the female and the male speaker resulting in a total of 120 trials.

RESULTS

First, comparisons against chance indicated that participants' overall accuracy of the presented stimuli was significantly above chance (see **Figure 3**) [Audio condition: $M = 84.1$, $SD = 9.2$: t-test against chance with equal variance not assumed $t_{(19)} = 24.7$, $p < 0.0001$; Match condition: $M = 79$, $SD = 8.8$, $t_{(19)} = 23.5$, $p < 0.0001$; Mismatch condition: $M = 69.1$, $SD = 5.2$, $t_{(19)} = 31$, $p < 0.0001$]. In order to determine participants' performance in each of the three conditions we calculated the F-score (2*accuracy*completeness)/(accuracy+completeness): the harmonic mean of Accuracy [#hits/(#hits+#false alarms)] and Completeness (#hits/(#hits+#misses)). We ran a repeated measures ANOVA with Condition (Audio Only, Audio-Gesture Match, Audio-Gesture Mismatch) and Type of Prosodic Contour (IP and PP) as within-subject factors. We found a significant main effect for condition [$F_{(2, 18)} = 20.1$, $p = 0.0001$, $\acute{\eta}^2 = 0.7$], a marginally significant effect for Type [$F_{(1, 19)} = 4.226$, $p = 0.054$, $\acute{\eta}^2 = 0.18$] and a significant interaction of Type and Condition [$F_{(2, 18)} = 14.624$, $p < 0.0001$, $\acute{\eta}^2 = 0.6$]. Paired sample t-tests used for *post-hoc* comparisons (Bonferroni correction $p < 0.0083$) revealed a significant difference between Audio Only ($M = 84.1$, $SD = 9.2$) and Audio-Gestures Mismatch ($M = 69.1$, $SD = 5.2$) conditions [$t_{(19)} = 6.78$, $p < 0.0001$], and between Audio-Gesture Match ($M = 79$, $SD = 8.8$) and Audio-Gesture Mismatch conditions [$t_{(19)} = 4.67$, $p < 0.0001$], but not between Audio only and Audio-Gesture Match conditions [$t_{(19)} = 1.40$, $p = 0.178$]. While the type of the prosodic contour did not affect participants' performance in the Audio only condition [$M_{IP} = 87$, $SD_{IP} = 10$; $M_{PP} = 79$, $SD_{PP} = 13$: $t_{(19)} = 2.408$, $p = 0.026$], participants performed significantly better on sentences disambiguated with PP than on sentences disambiguated with IP boundaries in Audio-Gesture Match [$M_{IP} = 75$,

$SD_{IP} = 11$; $M_{PP} = 85$, $SD_{PP} = 12$: $t_{(19)} = -3.105$, $p = 0.006$] and Audio-Gesture mismatch [$M_{IP} = 64$, $SD_{IP} = 8$; $M_{PP} = 70$, $SD_{PP} = 10$: $t_{(19)} = -3.376$, $p = 0.003$] conditions. First, these results show that matching gestures do not lead to a better comprehension than audio alone, while mismatching gestures hinder comprehension. Second, when the prosody of gestures mismatched that of speech, participants could not ignore the mismatch in their effort to disambiguate sentences. Interestingly, while on the whole, perceiving speech with and without gestures did not appear to influence sentence comprehension as scores are above chance level, participants have more difficulties to disambiguate sentences with IP than with PP boundaries both in the gestures matched and in the gestures mismatched conditions.

GENERAL DISCUSSION

Our findings show that when presented with acoustic linguistic stimuli that contain only prosodic information (i.e., low-pass filtered speech), participants are highly proficient in detecting whether speech sounds and gestures match. The prosodic information of spoken language must therefore be tightly connected to gestures in speech production that are exploited in speech perception. The syntactic structure and the meaning of utterances appear thus not to be necessary for the perceiver to align gestures and prosody. Additionally, participants could also use co-speech gestures in their comprehension of potentially ambiguous sentences, i.e., sentences with the same sequence of words, thus totally ambiguous in their written form, but with different prosodic structures. The disambiguation of these sentences could be triggered either by the PP or by the IP division into constituents. Our results show that matching gestures do not lead to a better comprehension than audio alone, while mismatching gestures led participants to choose significantly more the meaning signaled by gestures. Therefore, gestures are used in interpreting the meaning of ambiguous sentences. Interestingly, in the presence of gestures, participants have more difficulties to disambiguate sentences with IP than with PP boundaries in both conditions. These results suggest that the presence of gestures impairs performances when auditory cues are stronger. For example, it is possible that PPs are less marked by auditory cues than the IPs and therefore gestures might give additional information in this case. It seems also important here to point out the fact that in the present study what we call mismatch videos are videos in which the audio file of one meaning of a sentence is presented with the image video of the alternative meaning of the same sentence. Therefore, this manipulation of stimuli could have led to a possible artifact in the participants' performances. Though this possibility cannot be excluded entirely, we believe it is unlikely. At the end of the test session, we asked participants whether they had noticed the mismatching manipulation. None of the participants tested reported any perception of a manipulation. Thus, when they had the two categories of sentences, matched and mismatched, they did not detect that they were different because one was manipulated and not the other.

As opposed to the visual perception of speech in the speakers' face, where the movements of the mouth, the lips, but also the eyebrows (Krahmer and Swerts, 2004) are unavoidable in the production of spoken language, the gestures that accompany speech belong to a different category that is avoidable in speech production. Even though mismatching gestures decrease the intelligibility of spoken language, the addition of matching gestures does not appear to give an advantage over speech perception in the auditory modality alone. We are, in fact, able to understand the meaning of sentences when talking on the phone, or if our interlocutor is for other reasons invisible. Our results, however, suggest that the prosody of language extends from the auditory to the visual modality in speech perception.

This link between speech and gestures is congruent with neuropsychological evidence for a strong correlation between the severity of aphasia and the severity of impairment in gesturing (Cocks et al., 2013). While further studies are clearly needed to identify the specific aspects of spontaneous gestures that are coordinated with speech acts, our results demonstrate that part of speech perception includes the anticipation that bodily behaviors, such as gestures, be coordinated with speech acts. Prosodic Phonology thus appears—at least in part—not to be a property exclusive to oral language. In fact, it has abundantly been shown to characterize also sign languages where it has an influence on all body movements (Nespor and Sandler, 1999; Wilbur, 1999; Sandler, 2011, 2012). It is also—at least in part—not specific to language. Previous findings have shown that part of prosody, i.e., rhythmic alternation as defined by the Iambic—Trochaic Law (Bolton, 1894; Nespor et al., 2008; Bion et al., 2011) characterizes also the grouping of non-linguistic visual sequences (Peña et al., 2011). Thus, language is a multimodal experience and some of its characteristics are domain-general rather than domain-specific.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://www.frontiersin.org/journal/10.3389/fpsyg.2014.00700/abstract

Movie S1 | One sentence low-pass filtered speech, match condition (Experiment 1).

Movie S2 | Same sentence low-pass filtered speech, mismatch condition (Experiment 1).

Movie S3 | One sentence, match condition (Experiment 2).

Movie S4 | Same video, mismatch condition (Experiment 2).

REFERENCES

Alibali, M. W., Kita, S., and Young, A. J. (2000). Gesture and the process of speech production: we think, therefore we gesture. *Lang. Cogn. Process.* 15, 593–613. doi: 10.1080/016909600750040571

Beach, C. M. (1991). The interpretation of prosodic patterns at points of syntactic structure ambiguity: evidence for cue trading relations. *J. Mem. Lang.* 30, 644–663. doi: 10.1016/0749-596X(91)90030-N

Bernstein, L. E., Eberhardt, S. P., and Demorest, M. E. (1998). Single-channel vibrotactile supplements to visual perception of intonation and stress. *J. Acoust. Soc. Am.* 8, 397–405.

Bion, R. H., Benavides, S., and Nespor, M. (2011). Acoustic markers of prominence influence infants' and adults' memory for speech sequences. *Lang. Speech* 54, 123–140. doi: 10.1177/0023830910388018

Blossom, M., and Morgan, J. L. (2006). "Does the face say what the mouth says? A study of infants' sensitivity to visual prosody," in *Proceedings of the 30th Annual Boston University Conference on Language Development*, eds D. Bamman, T. Magnitskaia, and C. Zaller (Somerville, MA: Cascadilla Press).

Bolton, T. (1894). Rhythm. *Am. J. Psychol.* 6, 145–238. doi: 10.2307/1410948

Carlson, K., Clifton, C., and Frazier, L. (2001). Prosodic boundaries in adjunct attachment. *J. Mem. Lang.* 45, 58–81. doi: 10.1006/jmla.2000.2762

Christophe, A., Peperkamp, S., Pallier, C., Block, E., and Mehler, J. (2004). Phonological phrase boundaries constrain lexical access: I. Adult data. *J. Mem. Lang.* 51, 523–547. doi: 10.1016/j.jml.2004.07.001

Chu, M., and Kita, S. (2011). The nature of gestures' beneficial role in spatial problem solving. *J. Exp. Psychol. Gen.* 140, 102–115. doi: 10.1037/a0021790

Cocks, N., Dipper, L., Pritchard, M., and Morgan, G. (2013). The impact of impaired semantic knowledge on spontaneous iconic gesture production. *Aphasiology* 27, 1050–1069. doi: 10.1080/02687038.2013.770816

Cook, S. M., and Goldin-Meadow, S. (2006). The role of gesture in learning: Do children use their hands to change their minds? *J. Cogn. Dev.* 7, 211–232. doi: 10.1207/s15327647jcd0702_4

Cooper, W. E., and Paccia-Cooper, J. (1980). *Syntax and Speech.* Cambridge, MA: Harvard University Press. doi: 10.4159/harvard.9780674283947

Corballis, M. C. (2002). *From Hand to Mouth: the Origins of Language.* Princeton; Oxford: Princeton University Press.

Cutler, A., Dahan, D., and van Donselaar, W. (1997). Prosody in the comprehension of spoken language: a literature review. *Lang. Speech* 40, 141–201.

Dachkovsky, S., Healy, C., and Sandler, W. (2013). Visual intonation in two sign languages *Phonoly* 30, 211–252. doi: 10.1017/S0952675713000122

de Ruiter, J. P., Bangerter, A., and Dings, P. (2012). The interplay between gesture and speech in the production of referring expressions: investigating the tradeoff hypothesis. *Top. Cogn. Sci.* 4, 232–248. doi: 10.1111/j.1756-8765.2012.01183.x

Esposito, A., Esposito, D., Refice, M., Savino, M., and Shattuck-Hufnagel, S. (2007). "A preliminary investigation of the relationship between gestures and prosody in Italian," in *Fundamentals of Verbal and Nonverbal Communication and the Biometric Issue*, eds A. Esposito, M. Bratanic, E. Keller, and M. Marinaro (Amsterdam: IOS Press), 45–57.

Fernald, A., and McRoberts, G. W. (1995). "Prosodic bootstrapping: a critical analysis of the argument and the evidence," in *Signal to Syntax: Bootstrapping from Speech to Syntax in Early Acquisition,* eds J. L. Morgan and C. Demuth (Hillsdale, NJ: Erlbaum Associates), 365–387.

Graf, P. H., Cosatto, E., Strom, V., and Huang, F. J. (2002). "Visual prosody: Facial movements accompanying speech," in *Proceedings of the Fifth IEEE International Conference on Automatic Face and Gesture Recognition* (Washington, DC). doi: 10.1109/AFGR.2002.1004186

Guellaï, B., Coulon, M., and Streri, A. (2011). The role of motion and speech in face recognition at birth. *Vis. Cogn.* 19, 1212–1233. doi: 10.1080/13506285.2011.620578

Iverson, J. M., and Goldin-Meadow, S. (1998). Why people gesture when they speak. *Nature* 396, 228.

Kelly, S., and Barr, D. (1999). Offering a hand to pragmatic understanding: The role of speech and gesture in comprehension and memory. *J. Mem. Lang.* 40, 577–592.

Kendon, A. (1983). Gesture. *J. Vis. Verb. Lang.* 3, 21–36.

Kendon, A. (1994). Do gestures communicate? *Res. Lang. Soc. Inter.* 27, 175–200.

Kjelgaard, M. M., and Speer, S. R. (1999). Prosodic facilitation and interference in the resolution of temporary syntactic closure ambiguity. *J. Mem. Lang.* 40, 153–194. doi: 10.1006/jmla.1998.2620

Knoll, M. A., Uther, M., and Costall, A. (2009). Effects of low-pass filtering on the judgement of emotional affect using speech directed to infants, adults and foreigners. *Speech Commun.* 51, 210–216. doi: 10.1016/j.specom.2008.08.001

Krahmer, E. J., and Swerts, M. (2004). "More about brows: a cross-linguistic analysis-by-synthesis study," in *From Brows to Trust: Evaluating Embodied Conversational Agents*, eds C. Pelachaud and Z. S. Ruttkay (Antwerp: Kluwer Academic Publishers), 191–216.

Krahmer, E. J., and Swerts, M. (2007). The effects of visual beats on prosodic prominence: acoustic analyses, auditory perception and visual perception. *J. Mem. Lang.* 57, 396–414. doi: 10.1016/j.jml.2007.06.005

Kraljic, T., and Brennan, S. E. (2005). Prosodic disambiguation of syntactic structure: for the speaker or for the addressee? *Cogn. Psychol.* 50, 194–231. doi: 10.1016/j.cogpsych.2004.08.002

Ladd, D. R. (1986). Intonational phrasing: the case for recursive prosodic structure. *Phonol. Yearbook* 3, 311–340. doi: 10.1017/S0952675700000671

Langus, A., Marchetto, E., Bion, R. A., and Nespor, M. (2012). Can prosody be used to discover hierarchical structure in continuous speech? *J. Mem. Lang.* 66, 285–306. doi: 10.1016/j.jml.2011.09.004

Lehiste, I. (1973). Phonetic disambiguation of syntactic ambiguity. *Glossa* 7, 102–122.

Lehiste, I., Olive, J. P., and Streeter, L. (1976). Role of duration in disambiguating syntactically ambiguous sentences. *J. Acoust. Soc. Am.* 60, 1199–1202. doi: 10.1121/1.381180

Locke, J. L., Bekken, K. E., McMinn-Larsen, L., and Wein, D. (1995). Emergent control of manual and vocal-motor activity in relation to the development of speech. *Brain Lang.* 51, 498–508. doi: 10.1006/brln.1995.1073

Marslen-Wilson, W. D., Tyler, L. K., Warren, P., Grenier, P., and Lee, C. S. (1992). Prosodic effects in minimal attachment. *Q. J. Exp. Psychol.* 45, 73–87. doi: 10.1080/14640749208401316

McGurk, H., and MacDonald, J. (1976). Hearing lips and seeing voices. *Nature* 264, 746–748. doi: 10.1038/264746a0

McNeill, D. (1992). *Hand and Mind: What Gestures Reveal About Thought.* Chicago, IL: University of Chicago Press.

McNeill, N. M., Alibali, M. W., and Evans, J. L. (2000). The role of gesture in children's comprehension of spoken language: now they need it, now they don't. *J. Nonverbal Behav.* 24, 131–150. doi: 10.1023/A:1006657929803

Munhall, K. G., Jones, J. A., Callan, D. E., Kuratate, T., and VatikiotisBateson, E. (2004).Visual prosody and speech intelligibility. *Psychol. Sci.* 15, 133–137. doi: 10.1111/j.0963-7214.2004.01502010.x

Nagel, H. N., Shapiro, L. P., Tuller, B., and Nawy, R. (1996). Prosodic influences on the resolution of temporary ambiguity during on-line sentence processing. *J. Psycholinguist. Res.* 25, 319–344. doi: 10.1007/BF01708576

Nespor, M., Shukla, M., van de Vijver, C., Avesani, H., and Schraudolf, and, C., Donati (2008). Different phrasal prominence realization in VO and OV languages. *L&L* 7, 1–28.

Nespor, M., and Sandler, W. (1999). Prosody in israeli sign language. *Lang. Speech* 42, 143–176. doi: 10.1177/00238309990420020201

Nespor, M., and Vogel, I. (1986). *Prosodic Phonology.* Dordrecht: Foris, 327.

Nespor, M., and Vogel, I. (2007). *Prosodic Phonology.* 1st Edn. 1986. Berlin: Mouton De Gruyter.

Peña, M., Bion, R. H., and Nespor, M. (2011). How modality specific is the iambic-trochaic law? Evidence from vision. *J. Exp. Psychol. Lang. Mem. Cogn.* 37, 1199–1208. doi: 10.1037/a0023944

Pfau, R., and Quer, J. (2010). "Nonmanuals: their grammatical and prosodic roles," in *Sign Languages,* ed D. Brentani (Cambridge: Cambridge University Press), 381–402.

Ping, R., and Goldin-Meadow, S. (2010). Gesturing saves cognitive resources when talking about non-present objects. *Cogn. Sci.* 34, 602–619. doi: 10.1111/j.1551-6709.2010.01102.x

Price, P. J., Ostendorf, M., Shattuck-Hufnagel, S., and Fong, C. (1991). The use of prosody in syntactic disambiguation. *J. Acoust. Soc. Am.* 90, 2956–2970. doi: 10.1121/1.401770

Pynte, J., and Prieur, B. (1996). Prosodic breaks and attachment decisions in sentence parsing. *Lang. Cogn. Process.* 11, 165–192. doi: 10.1080/016909696387259

Rimé, B., and Shiaratura, L. (1991). "Gesture and speech," in *Fundamentals of Non-Verbal Behaviour,* eds R. S. Feldman and B. Rime (Cambridge: Cambridge University Press), 239–281.

Sandler, W. (1999). Prosody in two natural language modalities. *Lang. Speech* 42, 127–142. doi: 10.1177/00238309990420020101

Sandler, W. (2005). Prosodic constituency and intonation in sign language. *Linguistische Berichte* 13, 59–86.

Sandler, W. (2011). "The phonology of movement in sign language," in *Blackwell Companion to Phonology,* eds M. van Oostendorp, C. Ewen, K. Rice, and E, Hume (Oxford: Wiley-Blackwell), 577–603.

Sandler, W. (2012). Dedicated gestures the emergence of sign language. *Gesture* 12, 265–307. doi: 10.1075/gest.12.3.01san

Schafer, A., Speer, S., and Warren, P. (2005). "Prosodic influences on the production and comprehension of syntactic ambiguity in a game-based conversation task,"

in *Approaches to Studying World Situated Language Use,* eds M. Tanenhaus and J. Trueswell (Cambridge: MIT Press).

Selkirk, E. O. (1984). *Phonology and Syntax: the Relation Between Sound and Structure*. Cambridge: MIT Press.

Singer, M. A., and Goldin-Meadow, S. (2005). Children learn when their teachers gestures and speech differ. *Psychol. Sci.* 16, 85–89. doi: 10.1111/j.0956-7976.2005.00786.x

Snedeker, J., and Trueswell, J. C. (2003). Using prosody to avoid ambiguity: Effects of speaker awareness and referential context. *J. Mem. Lang.* 48, 103–130. doi: 10.1016/S0749-596X(02)00519-3

So, W. C., Kita, S., and Goldin-Meadow, S. (2009). Using the hands to identify who does what to whom: gesture and speech go hand-in-hand. *Cogn. Sci.* 33, 115–125. doi: 10.1111/j.1551-6709.2008.01006.x

Speer, S. R., and Blodgett, A. (2006). "Prosody," in *Handbook of Psycholinguistics, 2nd Edn.,* eds M. Traxler and M. A. Gernsbacher (Amsterdam: Elsevier), 505–537.

Speer, S. R., Warren, P., and Schafer, A. J. (2011). Situationally independent prosodic phrasing. *Lab. Phonol.* 2, 35–98. doi: 10.1515/labphon.2011.002

Sumby, W. H., and Pollack, I. (1954). Visual contribution to speech intelligibility in noise. *J. Acoust. Soc. Am.* 26, 212–215. doi: 10.1121/1.1907309

Warren, P. (1999). "Prosody and sentence processing," in *Language Processing,* eds S. Garrod and M. Pickering (Hove: Psychology Press), 155–188.

Warren, P., Grabe, E., and Nolan, F. (1995). Prosody, phonology and parsing in closure ambiguities. *Lang. Cogn. Process.* 10, 457–486. doi: 10.1080/01690969508407112

Weber, A., Grice, M., and Crocker, M. W. (2006). The role of prosody in the interpretation of structural ambiguities: a study of nticipatory eye movements. *Cognition* 99, B63–B72. doi: 10.1016/j.cognition.2005.07.001

Wilbur, R. B. (1999). Stress in ASL: Empirical evidence and linguistic issues. *Lang. Speech* 42, 229–250. doi: 10.1177/00238309990420020501

Yasinnik, Y., Renwick, M., and Shattuck-Hufnagel, S. (2004). "The timing of speech-accompanying gestures with respect to prosody," in *From Sound to Sense: 50+ Years of Discoveries in Speech Communication, 11–13 June 2004,* (Cambridge, MA).

Order of the major constituents in sign languages: implications for all language

Donna Jo Napoli[1]* and Rachel Sutton-Spence[2]

[1] Department of Linguistics, Swarthmore College, Swarthmore, PA, USA
[2] Post-Graduate Department of Translation, Federal University of Santa Catarina, Florianopolis, Brazil

A survey of reports of sign order from 42 sign languages leads to a handful of generalizations. Two accounts emerge, one amodal and the other modal. We argue that universal pressures are at work with respect to some generalizations, but that pressure from the visual modality is at work with respect to others. Together, these pressures conspire to make all sign languages order their major constituents SOV or SVO. This study leads us to the conclusion that the order of S with regard to verb phrase (VP) may be driven by sensorimotor system concerns that feed universal grammar.

Keywords: sign languages, word order, vision, syntax, sensorimotor systems and language, gesture

INTRODUCTION

In the initial period of linguistic analysis of sign languages, scholars tended to stay away from examining phenomena that were modality bound in favor of those that were more universal, in order to establish that sign languages were bona fide languages (see Woll, 2003 for an overview). Since the mid-1980s, however, scholars have turned their attention to the importance of modality (Bergman and Wallin, 1985; Sze, 2003; also see Meier et al., 2002).

We focus attention on the issue of sentence-level sign order in sign languages, looking at subject, object and verb. Research on 42 sign languages (see **Table 1**), taken as a whole, coupled with our own observations leads to generalizations about order that contrast to varying degrees with word order in spoken languages. We consider two hypotheses: (1) that our generalizations are due to universal pressures on language, ones which are seen most strongly in young languages, and (2) that our generalizations are due to modality; that is, the patterns for sign order in sign languages are determined by what makes sense visually. We conclude that the first hypothesis carries us quite far, but consideration of visual pressures allows us to account for all the observed tendencies in our study. We conclude that all sign languages should order their constituents SOV and SVO in most declaratives. Importantly, this does not preclude the possibility that languages may impose language-specific constraints on order within a phrase (see work on noun phrases in Estonian Sign Language, Miljan, 2000, and Taiwanese Sign Language, Zhang, 2007).

TERMINOLOGY REGARDING PREDICATES AND NOMINALS

We use V throughout to indicate predicates of any category. We use S and O to refer to the arguments of a V, but these labels are problematic, since what is referred to as S in the literature

is typically agent and what is referred to as O is typically any other argument. We do not include discussion of non-argument nominals.

As for nominals, to understand the generalizations here we must pay attention to articulation. Referents can be manually articulated via a lexical NP (including fingerspelling) or via finger pointing to an object within sight. These are two typical ways of introducing referents (what we here call players) into the discourse (what we here call the conversational scene).

Once a player is on the scene, it is commonly assigned a spatial index and subsequently this index is pointed to Johnston (2013). Many behaviors fall under the rubric "pointing": referential spatial indexes can be pointed to by finger, gaze, lip, chin, head-tilt, among others. Further, already introduced arguments can be incorporated into a V (Wilbur, 2003), or indicated by body shift (Bahan, 1996) and/or embodiment by the signer (Meir et al., 2007). For justification of including all these mechanisms as ways to indicate arguments, see Neidle et al. (2000). Still, null arguments are possible (Lillo-Martin, 1986). Where a sentence appears to have an "omitted" argument (i.e., no articulatory realization, manual or non-manual), we take such an argument to be expressed earlier in the discourse or to be understood through context, otherwise the sentence would be incomprehensible (Bergman and Wallin, 1985, p. 220). Argument omission is typical with a series of verbs that have the same argument (often S), where that argument has already been established (McIntire, 1980; Padden, 1988). Note that "I" and "you" are always on the scene, since they are participants in the sign/speech act.

Since ways of referring to old-information referents are, with one exception, layered (i.e., built into the V or indicated by the non-manuals), one cannot talk about their order with respect to the V: they are expressed simultaneously. We understand these

Table 1 | Sign Languages.

Adamorobe	Nyst, 2007
Al-Sayyid Bedouin	Sandler et al., 2005; Padden et al., 2010
American	Fischer, 1975, 1990; Friedman, 1976; Klima and Bellugi, 1979; Baker and Cokely, 1980; McIntire, 1980; Woodward, 1980; Liddell and Johnson, 1986; Lillo-Martin, 1986, 1991; Wilbur, 1987, 2002, 2003; Padden, 1988, 1990; Fischer and Janis, 1990; Kegl, 1990; Liddell, 1990; Petronio, 1993; Bahan, 1996; Matsuoka, 1997; Neidle et al., 2000; Taub, 2001; Aronoff et al., 2003; Chen Pichler et al., 2008
Argentine	Massone and Curiel, 2004
Australian	Johnston and Schembri, 2007a,b; Johnston et al., 2007
Austrian	Wilbur, 2002; Chen Pichler et al., 2008
Brazilian	Quadros, 2003; Quadros and Quer, 2008
British	Deuchar, 1983; Sutton-Spence and Woll, 1999
Catalan	Morales-López et al., 2005; Quadros and Quer, 2008
Chinese	Yau, 2008
Colombian	Oviedo, 2001
Croatian	Milkovic et al., 2006, 2007; Chen Pichler et al., 2008
Danish	Engberg-Pedersen, 1994; Kristoffersen, 2003
Estonian	Miljan, 2000
Finnish	Jantunen, 2008
Flemish	Vermeerbergen, 1996; Johnston et al., 2007; Vermeerbergen et al., 2007a,b
French	Baron, 1998; Cuxac, 2000; De Langhe et al., 2004; Cuxac and Sallandre, 2007
French Swiss	Boyes Braem et al., 1990
German	Glück and Pfau, 1998; Leuninger, 2000; Hänel, 2005; Happ and Verköper, 2006; Plaza-Pust, 2008; Plaza-Pust and Weinmeister, 2008
Greek	Efthimiou and Fotinea, 2007
Hong Kong	Sze, 2003
Indian	Zeshan, 2003; Aboh et al., 2005
Inuit	Schuit et al., 2011
Irish	Johnston et al., 2007
Israeli	Meir, 1995; Aronoff et al., 2003; Rosenstein, 2004; Meir and Sandler, 2007; Meir et al., 2007, 2010a,b
Italian	Volterra et al., 1984; Corrazza et al., 1985; Boyes Braem et al., 1990
Japanese	Nakanishi, 1994
Kenyan	Akach, 1992; Jefwa, 2009
Malagasy	Minoura, 2008
Mexican	Quinto, 1999
New Zealand	McKee and Kennedy, 2005
Polish	Farris, 1994; Wojda, 2010
Portuguese	Delgado-Martins et al., 1994
Providence Island Language	Washabaugh et al., 1978; Woodward, 1987
Quebec	Nadeau and Desouvrey, 1994
Russian	Kimmelman, 2012
The Netherlands	Bogaerde and Mills, 1994; Coerts, 1994
South African	Vermeerbergen et al., 2007a,b
Spanish	Bobillo-García et al., 2006; Morales-López et al., 2012
Swedish	Bergman and Wallin, 1985
Taiwanese	Smith, 2005; Zhang, 2007; Tsay and Myers, 2009
Turkish	Zeshan, 2006

referents in the context of the discourse and of our knowledge of who the signer/speaker is, and what the signer/speaker might be trying to communicate; this is a general practice in language comprehension (Carston, 2002, among others).

The only exception is manual (i.e., finger or hand) pointing; this is generally not simultaneously articulated with the V. Many of our sources do not indicate manual pointing, but we use any information they do present. We categorize lexical NPs and NPs indicated by manual pointing together under the rubric "manually-expressed NPs," and we use the abbreviation MNP.

A WORD ON DATA

We surveyed articles on 42 sign languages, as shown in **Table 1**, where language names are given in English (cited studies tell which varieties of language are gathered under these rubrics). We draw upon data collected and analyzed in these works as well as cite insights of others, without necessarily adopting the authors' analyses.

While some conclusions in these works seem resilient within the study of a given language and sometimes across languages, many are fragile in that they do not find corroboration in other studies. Brennan (1994) points out that American Sign Language, for example, has been analyzed as SVO (Fischer, 1975), V-final (Friedman, 1976), and topic-comment (Baker and Cokely, 1980). We add that ASL has been analyzed as varying between SVO and SOV depending on sociolinguistic factors (Woodward, 1980). Further, sometimes no constraints on word order emerge; in Malagasy Sign Language all possible permutations of S, O, and V occur (Minoura, 2008). Bouchard and Dubuisson (1995) and Bouchard (1996) argue that there is no base order in sign languages (and they say spoken language has this option, as well), looking at ASL and Quebec Sign Language.

Unfortunately, much of the confusion in the literature results from how the various studies were carried out. While replication of results is a revered principle in science, many times the best we can hope for is corroboration (Giles, 2006). But often not even corroboration is found on sign order. Johnston et al. (2007; see also Coerts, 1994) point out that attempts at comparing studies are confounded by the range of methodologies adopted in data collection, varying from elicitation based on drawings, to translations of sentences in a written language, to seeking grammaticality judgments of constructed sentences, to examining spontaneous or naturalistic data (monologs or dialogs).

Reliance on these methods, rather than on a large corpus of naturally occurring data gathered with no aim other than general linguistic study, is problematic (McEnery and Wilson, 1996). Such methods' reliability is even more doubtful for sign language study, where often the number of native signers consulted is small (Johnston and Schembri, 2007a). The sociolinguistics of Deaf communities complicates the issue further. Sign language communities are small minority communities whose language is young and without well-developed community-based standards of correctness and which have few true native signers (Johnston, 2013). Concerns about basing analyses of any language on very limited data and about what we can conclude from different methods of data collection abound (Sprouse, 2011; Weskott and Fanselow, 2011; Gibson and Fedorenko, 2013) and lead

to the conclusion that methodological options in accumulating evidence for syntactic analysis should be expanded.

With regard to sign order studies, Johnston et al. (2007) point out further that often information about the linguistic consultants that might be pertinent to language variation is not given, and that issues as fundamental as having consistent criteria (or even any explicit criteria) for what counts as a clause or a complete sentence remain unresolved (and see Crasborn, 2007; Jantunen, 2008). Here we take the relevant unit for discussion to be predicates and their constellations of arguments, regardless of repetition of various parts (as in V sandwiching/doubling, see Fischer and Janis, 1990; Kegl, 1990; Matsuoka, 1997). We take a light V and the main V it supports to be one predicate, an unproblematic analysis since no arguments intervene between the two in the data observed (as in signing GIVE plus HUG, rather than simply HUG—a rare construction, reported for Flemish Sign Language, but which might reveal spoken language influence, see Johnston et al., 2007).

The variety of theoretical approaches used, from syntactically based ones to semantically-pragmatically based ones, is another complicating factor (Johnston et al., 2007). Theoretical biases impose themselves in fundamental ways. Simply transcribing sign languages with a morpheme-by-morpheme gloss and then a translation into a spoken language can obscure the information (lexical and functional) in a sign and how it is packaged (Slobin, 2006); there is no way to represent linguistic data that is theory-neutral (Ochs, 1979). Thus, in any given study we may not know exactly what data are under consideration and, hence, exactly what we can conclude. Further, many of the findings in the various studies consist of generalizations often in the form of tables that give numbers of occurrences of templates such as OV, SOV, SVO, etc., but few actual examples, so that various comparisons we wanted to make were precluded. Given this lack of information, we have no choice but to transcribe sign streams in the way our sources do, rather than in a consistent transcription system that might be better suited for sign languages (such as the Berkeley Transcription System in Slobin, 2006). While inconsistent coding inhibits comparison, one advantage of using the form presented in our sources is that sometimes this form is given in the ambient spoken language, and thus may relate articulatory information, since mouthing is common (Crasborn et al., 2008).

Sign languages can allow variety in order for the same range of reasons spoken languages do, including stylistic and grammatical concerns. So the murky issue of a so-called unmarked word order arises (Leeson and Saeed, 2012). We have chosen to be inclusive for fear of excluding relevant data. Still, we restrict ourselves to declaratives (as do most works in our survey and as do studies of spoken languages). A handful of our sources focus on interrogatives, so that few examples from them are of use to us. Importantly, even when a study is on some issue other than sign order, the data presented support our claims here (as, for example, with Inuit Sign Language, in Schuit et al. (2011), where they explicitly set aside order as an issue they will not address).

Further, we are leery of relying on data not taken from spontaneous conversation, given confounding influences of the laboratory situation itself. This concern is of particular weight for sign languages since Deaf linguistic consultants can be influenced by perceived researchers' expectations based on grammatical

properties of the ambient spoken language (Deuchar, 1983, p. 76; Coerts, 1994). Nevertheless, we use data from all 42 languages regardless of how it was collected.

GENERALIZATIONS IN THE DATA

Here we list the generalizations we have found in the literature, augmented by our own observations of ASL and BSL conversations. These generalizations concern only MNPs, since all other nominals are expressed simultaneously to the V, precluding statements of linear ordering with respect to the V. So when we say S precedes V, we mean an S that is an MNP precedes the V, and so on. With the exception of the first, these generalizations are tendencies. The section A Comparison to Two Accounts discusses two accounts of these generalizations along with data that run counter to them.

GENERALIZATION ONE

SOV is grammatical in all sign languages.

Yau (2008) makes this claim and our survey confirms it. We offer a typical example from Finnish Sign Language (Jantunen, 2008, p. 99):

BOY APPLE BUY '(The) boy buys an apple.'

If there are three MNPs in the sentence (which is uncommon in conversation) and all are arguments, then all can precede the V, as in this example from Israeli Sign Language (Meir et al., 2010b, p. 276):

WOMAN BOX TABLE PUT-ON 'The woman puts the box on the table.'

GENERALIZATION TWO

If an argument affects the phonological shape of the V, it precedes V.

This includes classifier predicates (Emmorey, 2003), agreeing verbs (Wilbur, 1987; Liddell, 1990), pointing verbs (De Langhe et al., 2004), spatial verbs (Padden, 1988; and see Liddell, 1990), and argument-sensitive verbs (Klima and Bellugi, 1979; Volterra et al., 1984, p. 33; a.k.a. "imitating" predicates, in Vermeerbergen et al., 2007b). (All types of V in this paper are discussed in Padden, 1988, 1990; Quadros and Quer, 2008; Padden et al., 2010.) Only plain Vs are exempt. Evidence comes from explicit statements by scholars in the surveyed articles and our own observations.

Many studies exhibit only SOV sentences and explicitly claim that V must come finally. Others exhibit only SOV sentences but claim that the order is topic-comment (as in McKee and Kennedy, 2005, on New Zealand Sign Language). Others explicitly claim that if the V is a classifier, it must come finally, while still others say a classifier predicate usually comes finally.

Many studies note sentences with the structure SVOV, the template of V sandwiches, where the two Vs indicate the same action. Whether we have two clauses here or only one is a tricky matter, but not one we need to resolve. What matters for us is that the first V is typically a simple form, whereas the second shows variable phonological shape, sometimes with aspectual marking on it, but often with more iconic information than the basic form, some of which may be affected by the arguments. (Many have

noted for ASL that if a V is aspectually marked, its O precedes it even in single-V clauses, where the explanation involves raising the marked V to a right-branching functional projection, leaving the O in pre-verbal position, as in Chen Pichler, 2011; Fischer and Janis, 1990; Matsuoka, 1997; Braze, 2004.)

Here we see a V sandwich from Russian Sign Language where the second instance of the V is accompanied by a non-manual adverbial morpheme (Kimmelman, 2012, example 47):

face: doubtfully
LOOK G-R-U-Š-A   LOOK   '[He] looked at the pear doubtfully'

Several studies explicitly mention that agreeing Vs come in final position. In Brazilian Sign Language, SVO is argued to be the unmarked order (Quadros, 2003) but agreeing verbs can also come in final position, with SOV order (see also Quadros and Lillo-Martin, 2010). If pointing verbs are discussed at all, they are typically mixed into the discussion of agreeing verbs.

We turn now to argument-sensitive verbs. The studies we consulted that offer evidence about argument-sensitive verbs (whether they note it or not) show that MNPs precede argument-sensitive Vs. For example, Johnston et al. (2007) discuss sentences containing HUG in Irish Sign Language, Flemish Sign Language, and Auslan. Sometimes the first appearance of an argument of HUG is an MNP which follows the V, as in this example from Auslan (Johnston et al., 2007, p. 192):

BOY MEET HUG$_p$ GRANDMOTHER

We analyze the above as two clauses (as do the study authors), but significantly the first appearance of the O of HUG follows it (that is, GRANDMOTHER). And here the articulatory shape of HUG has not been adjusted to match the arguments. We indicate this fact with the subscript "p," showing this is a plain V. (Argument-sensitive Vs, unlike most agreeing Vs, only optionally incorporate their arguments.) However, a V sandwich example from Irish Sign Language has two instances of HUG, the first without phonological adjustment for the arguments (HUG$_p$) and the second with such adjustment (HUG$_s$, where the subscript "s" indicates this is an argument-sensitive realization of the V). We find that the MNPs representing the relevant arguments (the hugger and the hugged) precede the second instance of HUG (HUG$_s$) and, further, that the S precedes the O in this Irish Sign Language sentence (Johnston et al., 2007, p. 192):

BOY HUG$_p$ WITH OLD-GRANDMOTHER HUG$_s$

GENERALIZATION THREE
The most common sentence type has only one new argument, which precedes the V.

We offer a typical example from Indian Sign Language (Aboh et al., 2005, p. 22) in the completive aspect (COMPL):

YESTERDAY FATHER DIE COMPL 'Yesterday (my) father died.'

In fact, V S is generally unfound except when the V's sense introduces a player (which can be an event) onto the scene. Evidence for this generalization comes from explicit statements by scholars and our own observations. Additionally, we present evidence from so called split-sentence constructions.

Claims in the literature and our observations
First, sign languages usually express at most one MNP in a sentence, a fact some authors explicitly note. Many studies exhibit no V-initial sentences, again an observation often explicitly noted (and predicted in Minoura, 2008, p. 49, an idea proposed to her in personal correspondence by Susan Fischer). Other studies do have V-initial sentences, but the Vs function precisely to present or introduce a new argument, such as the existential verbs "seem," "exist," and the presentational verb "happen," as in this example from Kenyan Sign Language (Jefwa, 2009, p. 167):

HAPPEN ONE MZUNGU COME KENYA 'It happened one European came to Kenya.'

or possessives (some of which are presentational, see Kristoffersen, 2003; Johnston et al., 2007), as in this example from Swedish Sign Language (Bergman and Wallin, 1985, p. 219):

HAVE CAR I 'I have a car.'

Still, in Malagasy Sign Language several V-initial assertions with other types of verbs are reported, an example being (Minoura, 2008, p. 52 and ff.):

MANDRARAKA KAMIÖ VATO 'scatter truck rock' 'The truck scatters rocks.'

Minoura suggests the order in such examples is an influence from written Malagasy. (For remarks on the influence of written language order on sign order, see Fischer, 1975; Bogaerde and Mills, 1994; De Langhe et al., 2004; Milkovic et al., 2007; Yau, 2008; Wojda, 2010, who argues that this factor makes it impossible to determine the unmarked word order of Polish Sign Language.)

Split-sentence constructions
When one conveys a proposition in which the predicate has two arguments, and the referents of both are new to the conversation, a common tendency is to employ two clauses. The first introduces one MNP with a predicate that locates it or otherwise gives an identifying characteristic of it. That is, the first has a monadic V. The second clause introduces the other MNP with a dyadic V, that is, a V that takes two arguments. In the second clause the argument of the dyadic V that was introduced in the earlier clause is now not manually expressed.

In the first clause the MNP is the S of its clause per force. In the second clause, the MNP is typically the S. Very often, this second clause tells what the referent of the MNP in the second clause does to the referent of the MNP in the first clause. That is, the MNP in the first clause is coreferential with the O of the second clause (which is not manually expressed). This construction is known as "the split-sentence construction," and has been characterized as

$S_1V\ S_2V$, since each subject precedes its predicate, as exemplified here in Italian Sign Language (Volterra et al., 1984, p. 32):

BAMBINO SEDUTO MAMMA PETTINARE
child   seated   mother   comb
'The child sits and the mother combs (his) hair.'

This signing stream conveys that the mother combs the hair of the seated child. The point for us is that instead of signing this proposition in a single clause with two MNPs, the choice is to have two clauses with only one MNP per clause, where that MNP is the S of the predicate and precedes it.

GENERALIZATION FOUR

When two MNPs occur in a locational expression that forms a single clause, larger more immobile objects tend to precede smaller more mobile ones, regardless of theta role or grammatical function.

However, animacy complicates the situation (see remarks in the section Order and the Visual Modality). We are appealing here to properties of the referents of the signs, not to properties of the signs themselves.

This fact is explicitly remarked on by many, and it is subsumed under the figure-ground principle (Happ and Verköper, 2006). An example from German Sign Language is seen here (the example is from Leuninger, 2000, p. 238; the translation is from Plaza-Pust, 2008, p. 85):

WAND$_1$ JACKE ICH HANG_AN$_1$ 'I hang up the jacket on the
wall   jacket   I   hang-on   wall.'

GENERALIZATION FIVE

O is immediately adjacent to V.

Evidence for this comes from the order observed in the vast majority of examples in our survey. Certainly the order OSV occurs often in sign languages, but the literature overwhelmingly analyzes this as topicalization of O (indicated typically by prosodic cues and/or discourse contexts; Padden, 1988; Lillo-Martin, 1991; Petronio, 1993). This generalization supports the idea that there is a verb phrase (VP) in sign languages.

GENERALIZATION SIX

In reversible sentences with plain verbs, SVO is favored.

Several studies note this tendency, regardless of the word order a language exhibits in non-reversible sentences. This tendency is noted so often that when a language does not exhibit it, the authors typically explicitly say that (as for Sign Language of the Netherlands, Coerts, 1994). Surprisingly, a study of Flemish Sign Language found more variation in word order in reversible sentences (where we find SOV and OSV) than in non-reversible (where we find only SOV) (Vermeerbergen, 1996). For the languages that favor SVO with plain verbs in reversible sentences, it would seem that NP$_1$ V NP$_2$ order is not ambiguous (interpreted only as SVO), whereas NP$_1$ NP$_2$ V order is open to the readings SOV and OSV (and see Fischer, 1975). In contrast, Kimmelman (2012) points out for Russian Sign Language, that, since OSV is marked, the cues that go with topicalization of the O should eliminate ambiguity in reversible sentences. The observation captured

in generalization six remains, and we return to discussion of possible motivation in sections An Amodal Account and A Modal Account.

A COMPARISON TO TWO ACCOUNTS

We list the generalizations here for easy reference:

Generalization One. SOV is grammatical in all sign languages.
Generalization Two. If an argument affects the phonological shape of the V, it precedes V.
Generalization Three. The most common sentence type has only one new argument, which precedes the V.
Generalization Four. When two MNPs occur in a locational expression that forms a single clause, the larger more immobile objects tend to precede smaller more mobile ones, regardless of theta role or grammatical function.
Generalization Five. O is immediately adjacent to V.
Generalization Six. In reversible sentences with plain verbs, SVO is favored.

Taken together, we arrive at the generalization that SV is the order we find in most intransitive sign language sentences, and SOV and SVO are the orders for transitive sentences. Further, the choice between SOV and SVO is frequently determined by phonological considerations, where most of the time SOV should be preferred.

AN AMODAL ACCOUNT

One possible account of these generalizations is amodal: perhaps there are universal pressures on language that favor these patterns.

Consider generalization one. If we categorize languages by the six possible string permutations of S, O, and V, we find that together SOV and SVO characterize around 76% of spoken languages (Dryer, 2005), where SOV is dominant and SVO is not far behind. (For a current count, see Dryer's ongoing site http://wals.info/chapter/81. There, 41% of the 1377 spoken languages considered are SOV, and 35% are SVO.) Further, many V-initial languages also have an alternate word order with the S preceding the V, as in Arabic and Berber, in contrast to SOV languages, which tend to be strictly V-final in unmarked sentences (Tomlin, 1986; Herring and Paolillo, 1995; among many). We might therefore want to conclude that SOV or SVO is possible in all languages. The biggest problem for this conclusion is the Celtic family. Celtic languages have been claimed to be rigidly VSO except for main clauses in Breton and Cornish (Tallerman, 1998). There is not complete agreement on this, however. A drift toward SVO has been documented for Breton and Welsh (Raney, 1984; but see Willis, 1998 for Welsh), and a claim made that SVO is more frequent in modern Breton than VSO (Varin, 1979; but see Timm, 1989). We conclude that, on the whole, languages in general favor SOV, not just sign languages, and languages in general favor adjacency of V and O.

But the tendency for SOV is stronger in sign languages. Why? Some linguists argue that SOV is the default order for human language (including Givón, 1979; Newmeyer, 2000a). Newmeyer (2000b), in fact, claims SOV was the order in proto-language. Sign languages are young, so perhaps the acceptability of SOV in all sign languages follows. Indeed, could all the generalizations

we noted in the immediately preceding section hold of young languages in general?

Many languages are known to have changed diachronically from SOV to SVO. In Indo-European, this is the case with English (Canale, 1978, among many), Greek (Taylor, 1994), Swedish (Delsing, 2000), Icelandic (Hróarsdóttir, 2000: p. 60), Norwegian (Sundquist, 2006), Spanish (Parodi, 1995), and Italian (Antinucci et al., 1979). (And see Fischer, 2010 for discussion of word order change in general, with a focus on Indo-European languages.) In Sino-Tibetan, this is the case with Bai, the Karen languages of Thailand and Burma, and may be responsible for a number of complex word order facts in languages of China (Dryer, 2003). In Niger-Congo, this is the case with Bantu languages (Givón, 1975). And the list continues. Rarely, however, do we find diachronic change in the opposite direction (Gell-Mann and Ruhlen, 2011). Some exceptions are the Austronesian language Motu (Crowley, 1992), the Western Oceanic language Takia (Ross, 2001), the Tai-Kadai language Kamti Tai (Khanittanan, 1986), and a few others, where that change is argued to be an influence from contact with an SOV language. (For overview and citations see Van Gelderen, 2011). Further, emerging sign languages favor SOV strongly (Meir et al., 2010b).

With respect to generalization three, while there is an enormous literature on (in)transitivity, trying to estimate the prevalence of different valencies is far from obvious (as in Brew and Schulte im Walde, 2002). In the substantial literature on creoles, no one, to our knowledge, discusses the relative prevalence of intransitive to transitive sentences (see, for example, McWhorter, 2000). And we are aware of no literature on any spoken language that claims that a particular language or language family has a tendency toward having only one fully referential NP (that is, an NP that is not a pronoun or an anaphor) in a clause or about young languages having such a tendency.

With respect to generalization four and spoken languages, again there is considerable literature on locational, existential, and possessive expressions, which have a number of semantic similarities. But much of that literature is concerned precisely with those semantic properties (for example, Hoekstra and Mulder, 1990). Some of the literature, however, addresses word order. Clark (1978, p. 88), for example, notes that "roughly speaking" definite NPs precede indefinite ones in English and French sentences of this type. However, we know of no claims to the effect that the size or mobility of the referent of an NP matters in the determination of word order in spoken languages.

One can also look to word order in spoken creoles with respect to the claim that young languages favor adjacency of V and O—that is, to support the claim that generalization five is true of young languages, since creoles are by and large young languages. DeGraff (2003, 2005) surveys a number of creoles and shows that, despite claims to the contrary (as in Bickerton, 1981, 1990 and following), creoles are not an exceptional kind of language morphologically and syntactically. In particular, SVO is not the (near) universal word order for creoles. Instead, creole VPs can be OV or VO. Still, it appears that many more creoles are SVO than SOV (Julien, 2002). So the evidence from creoles is not compelling with respect to the claim that young languages favor SOV (generalization one).

With respect to generalization six, while many languages allow a wide range of ambiguities, word order can be sensitive to situations of potential ambiguity with regard to grammatical functions (particularly S and O); indeed, sometimes in potential-ambiguity contexts in spoken language we do not find the otherwise expected word orders (Craig, 1977 for Jacaltec, Kuhn, 2001 for German, Lee, 2001 for Hindi and Korean, Vulanoviæ, 2005 and Flack Potts, 2007 for Japanese). Speakers of English adjust their word order to avoid ambiguity when the visual context is the source of the potential ambiguity (Haywood et al., 2005). While we have found no mention that this tendency is stronger in young languages, it certainly appears to be evidence of a natural language principle.

The only remaining generalization to be addressed with respect to spoken languages, number two, calls for a more complex discussion. The situation in spoken languages is interestingly complex, and we restrict the discussion here to the tense-carrying V (not to participles, which enter into a different paradigm). In general, for an argument to affect the phonological shape of the V (an effect that is arbitrary with respect to meaning for spoken languages—we return to this point in the section Order and the Visual Modality, when we discuss generalization two), there must be agreement between the two. Most commonly, if there is agreement, the V agrees with the S alone. Since S precedes V in most languages, this is no problem for our generalization. However, nearly 9% of spoken languages are V-initial (conflating the VSO and VOS examples on the site http://wals.info/chapter/81), among them the Celtic languages. In all the Celtic languages, the V does not agree with an S that is a fully referential NP, but it might agree with a pronominal S (whether overt or "pro"), as happens in Welsh (Borsley and Roberts, 2005, p. 40). But the very conditions for a pronominal S are that the referent already be present on the conversational scene. This is consistent with our motivation for generalization two. On the other hand, various varieties of Arabic allow both VSO and SVO order, but the V still agrees with the S even when the S follows the V, although interesting complications arise. In particular, in Standard Arabic (as opposed to Lebanese or Moroccan Arabic) when the S follows the V we find agreement for gender only, not for person and number, but when the S precedes the V, we find agreement for the full range of features (Aoun et al., 1994; Alexiadou and Anagnostopoulou, 1998).

Further, some languages allow agreement of a V with O, either direct object (as with Hungarian, Ge'ez, and Eastern Aramaic) or both direct and indirect object (as with Amharic, Swahili, and Lebanese Arabic), where O might well-follow V. Again, we find interesting complications. In Lebanese Arabic, where O follows the V, V can agree with an O only if it is definite (Koutsoudas, 1969). The same is true of Swahili (Givón, 1976). Since definite NPs are used when the referent is already on the conversational scene, generalization two seems to loom in the background again. On the other hand, in Amharic a definite O triggers agreement on the V, while an indefinite O does not (Baker, 2012), going exactly counter to our expectations if generalization two holds of spoken languages.

We have not done a survey of agreement facts in general, and agreement is remarkably messy (see Moravcsik, 1988). However,

it seems clear that generalization two is not true of spoken languages, young or not, especially since we have found no typologists' claims to this effect.

In sum, an amodal account explains the preference for SOV, for the adjacency of V and O, and for word order to resolve potential ambiguities that arise in reversible sentences. But it does not account for the preference for clauses with only one fully referential NP, for word order in existentials and presentational sentences, nor for the phonological and semantic factors that affect word order in sign languages (i.e., generalizations two through four).

A MODAL ACCOUNT

The alternative account we now consider is that these generalizations are a result of the modality of sign languages.

With respect to generalization one, a number of studies of gesture conclude that SOV is the default order in visual communication. In one study, Gershkoff-Stowe and Goldin-Meadow (2002) had English speakers describe scenes solely with gesture, and in another they presented speakers with pictures and asked them to order them in a way that would communicate a given scene. In both, people presented scenes in the order SMA—stationary entity, moving entity, action. Importantly, the order of stationary before mobile entity is exactly what we find in sign languages, expressed in generalization four.

So et al. (2005) asked English speakers to describe vignettes in speech accompanied by gestures created on the spot as well as solely in gestures. When using gestures alone, the hands exploited space for reference and coreference more often than when speech was also used, and the types of entities the gestures represented differed. Most gestures accompanying speech concerned action, but gestures alone also concerned entities. From the data given, it appears that the order of "constituents" in gesture-only propositions resembles that in sign languages. For example, this is the description of a man communicating "man gives woman basket" with gestures (So et al., 2005, p. 1032):

> He first set up one person (man) on his body [G1] and a second person (woman) on his right [G2]. He then produced a GIVE gesture moving from a location in front of him (later identified as basket) to the location to his right (woman) [G3, which was coreferential with G2]. After producing a gesture for basket in the location in front of him [G4, which was coreferential with G3], he again produced a GIVE gesture moving from the basket location to the woman location [G5, which was coreferential with G2, G3, and G4].

We see clearly the strategy of setting up participants in an action, then expressing the action. And, when relevant, we see the strategy of setting up the S before other participants. Importantly, we see that the action gesture, whose articulatory shape is affected by the participants, appears after those participants, just as in sign languages (see generalization two).

Goldin-Meadow et al. (2008) likewise find that SOV recurs in non-verbal communication. They had native speakers of languages with varying word orders (English, Turkish, Spanish, Chinese) perform studies like those in Gershkoff-Stowe and Goldin-Meadow (2002)—using wholly gestures in one study

and arranging pictures in another, but now the scenes involved actions from an agent onto a patient (like transitive verbs) rather than intransitive changing-location actions. The order of constituents in speakers' native languages did not influence the order in these visual tasks. They conclude that SOV is the "natural order that we impose on events when describing and reconstructing them non-verbally and exploit when constructing language anew" (Goldin-Meadow et al., 2008, p. 9163).

Langus and Nespor (2010) replicated Goldin-Meadow et al.s' (2008) experiments with speakers of Italian and Turkish. Their results led them to a similar conclusion about the early stages of an emerging language: SOV is the preferred order in "simple improvised communication" (Langus and Nespor, 2010, p. 293). In another experiment they concluded that improvised communication does not organize its constituents hierarchically, in contrast to natural language. In a third experiment, they tested speech comprehension of sentences with prosodically flat words, where S, O, and V appeared in all possible orders and concluded that, while speakers understand best sentences whose order conforms to that of their native language (SVO for Italian; SOV for Turkish), compared reaction time in recognition of the meaning of speech strings with varying order shows a preference for V to precede O. They conclude that the computational system of grammar prefers SVO, whereas the preference for SOV in improvisational communication demonstrates "a direct link between the sensory-motor and the conceptual systems that prevails in gesture production" (Langus and Nespor, 2010, p. 308). In other words, SVO is the preferred syntactic order, with SOV being the natural conceptual order.

Gibson et al. (2013), in a gesture-production task with speakers of English, Japanese, and Korean, conclude that SOV is, indeed, the preferred order in gestural communication, but SVO arises when communication needs demand it, as in reversible events. The same is true in emerging sign languages; when asked to use gesture to describe reversible events in which both participants are animate ("girl kicks fireman"/"fireman kicks girl"), people prefer SVO (Meir et al., 2010a). This echoes the behavior of many sign languages, as stated in generalization six.

Gibson and colleagues tie this to works on language proper that claim SOV is the default order for human language. Their explanation for this shift to SVO in reversible events is based on the "noisy-channel" hypothesis (Shannon, 1948; Levy, 2008; Levy and Jaeger, 2007; the quote here is from Gibson et al., 2013, p. 1081).

> A speaker wishes to convey a meaning m and chooses an utterance u to do so. This utterance is conveyed across a channel that may corrupt u in some way, resulting in a received utterance ū. The noise may result from errors on the side of the producer, external noise, or errors on the side of the listener. The listener must use ū to determine the intended meaning m. The best strategy for a speaker is thus to choose an utterance u that will maximize the listener's ability to recover the meaning given the noise process.

Languages need to be robust against this omnipresent noise. Essentially, a representation of an event with an animate patient is more robust to noise when the agent and patient are separated by

the action (V). Spoken languages with SOV order can be robust against interfering noise by using case-marking, and they point out that case-marking is prevalent in SOV languages but almost absent in SVO languages.

Since languages are known to have changed diachronically from SOV to SVO, as discussed in section An Amodal Account, the idea that a noisy channel might be the impetus for such change arises. Hall et al. (2013) address this issue; they asked speakers of English to describe in pantomime both reversible and non-reversible transitive events. Critically, speakers always took on the role of actor, and Hall and colleagues noted what they call a "role conflict" in reversible events (Hall et al., 2013, p. 5):

> To describe a non-reversible event (e.g., a woman lifting a box) using SOV order, participants would generally adopt the role of the agent (long hair), then produce a gesture for the box without adopting any role. In this case, the participant does not need to do anything special to re-inhabit the role of agent in time to produce the action gesture. In contrast, using SOV for reversible events (e.g., a man lifting a woman) is likely to entail a role conflict between O and V. For example, if a participant described a reversible event using SOV order, she or he would first adopt the role of the agent (flexing muscles), then the patient (long hair). The participant is now in the patient role but is ready to produce the action, which requires him or her to be in the agent role. If the participant were to produce an action gesture without first doing something to switch back into the agent role, it may "feel" to him or her as if it is the patient and not the agent that is carrying out the action. It is this that we refer to as role conflict.

They suggest that the preference for SVO in reversible events is due to a desire to avoid role conflict. And they note that when speakers do produce SOV order in reversible events, they find ways to get around the potential role conflict, either by not embodying the role of the patient (perhaps simply tracing it in space) or by establishing a spatial location for agent and another for patient and then shifting appropriately between them when they pantomime the action. (Spatial marking is also observed in Gibson et al., 2013, who compare it to case marking in spoken languages.)

Schouwstra (2012) also addresses the issue of a natural word order by looking at gesture in an improvised communication experiment. Many of her findings echo those of earlier scholars. Her work differs, however, in arguing that constituent ordering is influenced not only by the cognitive abilities involved in making an analogy between language meaning and cognitive representations (and see de Swart, 2009), but also by the communicative needs involved in public expression, where the conventional nature of language imposes itself (Roberge, 2009). Participants view an event on a screen. Then they use gesture to describe it. The process of transitioning from the simultaneity of the picture to the linearization of the gesture string forces participants to consciously choose the order in which they present things. This choice can be made on grounds of communicative needs. Schouwstra makes a distinction between "motion events," which involve extensional predicates (that create transparent contexts), such as *carries* in "princess carries vase," and "intensional events," which involve intensional predicates

(that create opaque contexts), such as *think of* in "cook thinks of sock". Both Turkish and Dutch speakers strongly preferred SOV order in their gestural representation of motion events, but SVO order (though less strongly) in their gestural representation of intensional events. Schouwstra then looked at order in events involving a subset of intensional predicates, the creation verbs. She found that the tendency toward SVO was smaller with creation verbs than with other intensional verbs but was still the preferred order. (Indeed, we found evidence of pressure toward SVO with creation verbs in our study, but nothing conclusive.) There is no doubt that semantics influences word order in these experiments. As Schouwstra (2012, p. 148) says, "When making a sequence of the different elements, they [the participants] are forced to impose an order on the information. So it is only in making the information public, in being involved in communication, that ordering plays a role." Likewise, she found that when people interpret gesturing of others, SOV strings are interpreted more often as motion events than SVO strings are, and SVO strings are interpreted more often as intensional events than are SOV strings. "This shows that in emerging communication systems, meaning and structure have more to do with each other than previously thought. Moreover, it suggests that ordering information in utterances in these systems is quite an active process, rather than simply a reproduction of how information is represented mentally" (Schouwstra, 2012, p. 148).

Christensen and Tylén (2013) offer another gestural communication experiment which uses an interactional paradigm instead of an elicitation task. Participants communicate to a passive experimenter or a camcorder, thus participating in proper bidirectional communication, where dyads are dependent on mutual comprehension of the gestural systems that evolve during the experiment sessions. They followed up on Schouwstra's work, contrasting "object manipulation events" to "construction events," the latter of which involve effective verbs. The former consistently yielded SOV order, while the latter yielded SVO order, as we also found for sign languages, but with far too few examples to base a generalization upon. Again, we see that event structure rather than a cognitively natural order influences order in these gestural strings.

So the data on gestural communication is consistent with all the generalizations of section Generalizations in the Data.

Further, homesigners often produce strings of V plus one argument, where they place the V finally (that is, SV or OV) (Goldin-Meadow, 2003). And studies of young sign languages, still with a relatively unstable grammar, reveal a tendency for utterances to consist of SV, OV, and SOV (Senghas et al., 1997; Sandler et al., 2005; Haviland, 2011). These findings are, so far as they go, consistent with the generalizations of section Generalizations in the Data.

CONCLUSION

The amodal account covers some of our generalizations; the modal account covers all. One might then conclude that our observations on sign languages are evidence of a natural visual order. That is, we know vision is at play in both producing and receiving gestural strings and sign languages, so if one is to claim

some other cognitive ability is at play, the burden of proof lies on them.

Nevertheless, the fact that visual communication (gesture and sign languages) and spoken languages, particularly young languages, share important tendencies in order of constituents should make us wary of such a conclusion. It seems unlikely that totally independent pressures on sign languages and spoken languages would happen to produce such similarities. Two logical possibilities come up. One is that the pressures evidenced in the generalizations about order in sign languages really do hold of language in general, but that over time evidence for several of them has been lost as these pressures yield to competing pressures (whatever they might be), or several of them are simply gapped in spoken language. This possibility is not open to testing, unfortunately, but the speculation remains (and see Hale, 1975 for discussion of gaps in universals).

The other possibility is that the word order generalizations for sign languages reveal universal pressures augmented by visual pressures. As Chomsky (2013, p. 35) says, "… each language incorporates a mechanism that determines an infinite array of hierarchically structured expressions that are transferred for interpretation to two interfaces: the sensorymotor system SM for externalization and the conceptual-intentional system CI, for thought (broadly understood)." The structured expressions in spoken and sign languages are transferred to different sensorimotor systems—leading to different realizations.

At this point one might be led to the reasonable position that the universal pressures on word order discussed in this paper are grammatical in nature, while the pressures that apply only to sign language word order are visual in nature. Still, there is a way to see a coherence in the two sets of pressures. If, in fact, pressures of both the auditory and visual systems are behind the universal pressures on word order, we can view the sensorimotor pressures as motivating this particular part of universal grammar, which is apparent in both spoken and sign languages. Certainly, biological sources as a foundation for universal grammar should be seriously examined. After all, the innate language faculty, which serves for both spoken and sign languages, evolved somehow.

Given that language is embedded in the neuronal circuitry of the brain, and given that motor, cognitive, and perceptual systems are implicated in language learning and language use, we may assume that the language faculty should have come from pre-existing competencies, which initially were unrelated to language (Cowie, 2008; and, for compatible remarks, see even non-nativists, such as Tomasello, 2003). Certainly, finding evidence today that bears on human cognitive evolution is a daunting job, but our findings here suggest that comparative studies of languages in different modalities may offer new ways to approach the issue (and see Napoli and Sutton-Spence, 2011). Whatever the truth about language evolution may turn out to be, the birth of the language faculty will have been complex and, if we are correct, will involve many other competencies that developed earlier and were then adapted to language, with the sensorimotor systems playing a significant role.

The idea that shared language properties may follow from shared pressures of the visual and auditory sensorimotor systems seems to be gaining strength in the neuroscience field.

Tettamanti and colleagues argue (2005, p. 273), "… listening to sentences that describe actions engages the visuomotor circuits which subserve action execution and observation" (but see Mortan Ann Gernsbacher's remarks in Gallese et al., 2011). Further, the prevalence of SOV and SVO accords well with the representation of action in Broca's area (Kemmerer, 2012; but for arguments that Broca's area does not have a unified function, see Fedorenko et al., 2012). Additionally, neural tissue involved in language processing involves polymodal neural activity, so the idea that the different sensorimotor systems would share properties may follow from a cooperation of these neural activities (Petitto et al., 2000). And, finally, there is evidence that intellectual and perceptual-motor skills involve hierarchical unpacking of chunks of knowledge (Rosenbaum et al., 2001; Rosenbaum, 2009; Clark, 2013), thus sensorimotor-system pressures may even motivate the hierarchical nature of universal grammar.

Further, if this sensorimotor hypothesis about word order can be supported, it is the more interesting one since it calls for a reassessment of how to approach the issues of the order of the major constituents in language in general. Let us assume that the grammar of all languages organizes words into phrases, including VP. That means that OV and VO are both generated, depending on whether phrases in the language are head-initial or head-final. So the potential orders we can expect the relevant sensorimotor systems to produce in both spoken and sign languages for transitive sentences are SOV, OVS, SVO, and VOS. The fact that SOV and SVO occur so frequently in spoken language and so overwhelmingly in sign languages suggests that pressures of the sensorimotor systems favor S preceding VP. This accounts for the infrequency of spoken languages with unmarked word order being OVS (under 0.8%; 11 out of 1377) and VOS (under 2%; 25 out of 1377); they are bucking the sensorimotor system pressures. This also leads to the conclusion that OSV will be the result of topicalization from either SOV or SVO. That is, OSV should be a marked order in language, calling for contexts in which we are somehow highlighting the O. In fact, only 4 spoken languages out of 1377 have been claimed to have OSV as unmarked order (under 0.3%).

Finally, let's consider VSO. An immediate problem is that V and O are not adjacent. Further, we see no evidence of pressure from sensorimotor systems to favor V in initial position. As we discussed, VSO is (almost) non-existent in sign languages and is rare as an unmarked order in spoken languages (under 7%, 95 out of 1377). Importantly, as also discussed, VSO in spoken languages often has SVO as an alternate order. The strong consensus in the literature is that VSO arises from SVO via V-raising in order to satisfy requirements of the grammar (Choe, 1987; Carnie and Guilfoyle, 2000; for example), even for Irish (Bobaljik and Carnie, 1992). (For details on the analysis, see Alexiadou and Anagnostopoulou, 1998).

The sensorimotor hypothesis, then, says that S precedes VP as a fundamental strategy in language. This conclusion finds support in the language of people who are linguistically deprived in the sense that they were not exposed to accessible language during the early years of life. Such people generally manage to use appropriate word order in most situations, whereas many

other properties of language are problematic for them. This is true of Genie, an abused girl who was not rescued until the age of around thirteen (Curtiss et al., 1973; Fromkin et al., 1974; Curtiss, 1977; Goldin-Meadow, 1978) and of deaf "late learners" (Johnson and Newport, 1989; Newport, 1990; Newport et al., 2001; Wood, 2010). In fact, deaf children first exposed to ASL after the age of 6 do not produce appropriate variations in word order that native signers produce (even as young as 2 year olds), instead using SVO heavily (Lillo-Martin and Berk, 2003). That it is SVO rather than SOV that these late learners produce is consistent with the fact that their morphology is underdeveloped, thus their verbs exhibit fewer instances of phonological shape affected by arguments (that is, fewer instances of the situations that call for SOV, see discussion in section Generalization Two) than verbs of native signers (Newport, 1991). Thus, the sensorimotor hypothesis accounts for why some characteristics of language are "resilient" and others are "fragile" (Wood, 2007, 2010); the resilient ones are dependent upon sensorimotor pressures that exist regardless of language and that motivate certain parts of the grammar, while the fragile ones are not. In other words, late learners look at the world visually and their language is sensitive to visual pressures. On the other hand, they have trouble producing those grammatical structures that are not motivated by sensorimotor pressures, but are arbitrary to the particular language.

Given this explanatory force of the sensorimotor hypothesis, it is worth taking a closer look at what some of these pressures might be. The sensorimotor account of word order amounts to saying there are universal pressures driving the order similarities among sign and spoken languages, pressures that are imposed by factors that the visual and auditory sensorimotor systems have in common, and there are modality-specific pressures resulting in the order differences between sign and spoken languages, pressures imposed by the visual (-manual) sensorimotor system. In the next section we explore the relevant visual pressures on sign languages, and one suggestion of a pressure imposed by the manual articulators.

ORDER AND THE VISUAL MODALITY

Here we consider the generalizations that hold of sign languages but not of (young) spoken languages (i.e., generalizations two through five), and we argue each follows from visual needs or principles. Some of our accounts rely on coherence and iconicity; they turn upon the construction of a visual image, making testable predictions. There is no doubt that iconicity plays a role in sign language order. As De Langhe et al. (2004, p. 117) say (in our own translation), "... the most important thing for constructing sign expressions is iconicity... one must find the image that represents the subject and as soon as an image is formed in the mind, the translation into sign language becomes clear and easier." Thus, there is pressure for temporal and spatial organization to work together coherently at every level of grammar, maximizing comprehensibility.

GENERALIZATION TWO

If an argument affects the phonological shape of the V, it precedes V.

Why should sign languages but not spoken languages share this generalization? In a spoken language, the relationship between phonological features and meaning is (to a huge extent) arbitrary. In a sign language, that relationship is not arbitrary. Instead, the phonological shape of classifier predicates, agreeing verbs, pointing verbs, and argument-sensitive verbs will vary in non-arbitrary ways according to meaningful properties of their arguments, such as their spatial index and their size, shape, and general category (human, animal, small round object, and so on). For example, agreeing Vs involve a transfer of something (abstract or concrete) from one location to another. If visual perceptibility matters to the order of arguments, then we might expect an alignment such that the visual representation of transfer should involve a path that moves in the direction of the transfer. That is, the spatial-temporal organization should be coherent with the semantics of the utterance. This means that the point of initiation of the movement should be spatially indexed with the argument that is the origin of the transfer, and the endpoint of the movement should be spatially indexed with the argument that is the goal of the transfer (Meir, 1995; Aronoff et al., 2003). In most of the sign languages we have read about with respect to Vs of giving and taking, the verb GIVE moves from a point indexed with the giver to a point indexed with the receiver of the gift; whereas the verb TAKE moves from a point indexed with the one (or the place) from whom something is taken to a point indexed with the taker. In such examples as in classifier constructions, we find "mappings of envisioned mental spaces onto signing space" (Taub, 2001, p. 163). If the addressee is to make sense out of the phonological shape of these predicates, including the direction of path movement, the relevant arguments should already be present in the discourse or be introduced within the sentence before the V (for a similar claim, but with more conditions on it, see Yau, 2008, pp. 152–153).

With respect to classifier constructions, the non-arbitrariness of phonological features is rampant. To express that someone almost gave something to someone else, one might move only halfway along the path from one spatial index to another, for example (Quadros and Quer, 2008), perhaps with a dynamics that portrays hesitancy. Thus, iconicity can be a motivation with respect to the order of elements and with respect to various factors of a predicate's movement (direction, length of path, and so on), as well as with respect to other phonological parameters (such as facing of the hands, as in Meir, 2002). Syntactic structure is here a linguistic construction that itself conveys meaning (Goldberg, 1995, 2003).

As final evidence that generalization two reflects semantic concerns that are realized visually, we note that sign languages, like spoken languages, can exhibit phonological feature-spreading rules (as in compounding in ASL, see Liddell and Johnson, 1986). Such rules are purely phonological; they are arbitrary with respect to semantics, and in these instances feature spreading can be anticipatory as well as perseverative. So when phonological shape is arbitrary, signs can be affected by what follows linearly. It is only when phonological shape is meaningful (as with classifier predicates, agreeing verbs, pointing verbs, and argument-sensitive verbs) that the element that influences the phonological shape appears beforehand as the unmarked order.

Certainly it is possible to articulate a predicate whose phonological shape is affected by an argument before introducing the relevant argument (Padden, 1988), but this order is marked. The effect, according to the linguistic consultants we have asked, is like holding back information for dramatic impact and then revealing it, as in *And in walked. . .. her husband!*

An example from Inuit Sign Language makes the point nicely (Schuit et al., 2011, p. 21):

INDEX-LOC$_{3a}$ SCOOP DRILL-HOLE-WITH-AUGER FINISH. $_{3a}$WALK$_1$ TAKE-LONG-ITEM $_1$WALK$_{3a}$ WHITE-MAN CHISEL$_V$. DROP LONG-THIN-OBJECT MOVES-BELOW-SURFACE.
'Over there they started a hole with a scoop, and then drilled it with an ice-auger. Someone walked from there towards me and took my chisel. The white man walked back (to the hole) and used the chisel. Then he dropped it, and it went all the way to the bottom (of the sea).'

(The translation is Schuit et al.'s, but the following comments are ours.) In the second sentence, "$_{3a}$WALK$_1$" indicates that someone walked from spatial location 3a (where the scoop and then the ice-auger were used) to spatial location 1 (which is the signer's location). "TAKE-LONG-ITEM" indicates a classifier predicate in which someone is taking hold of a long item. "$_1$WALK$_{3a}$" indicates that someone walked from spatial location 1 back to spatial location 3a. And only now are we told that the someone was a white man and that the long object he took was the signer's chisel. Here an unspecified NP is spatially indexed; we can't see who it is, all we see is that he picked up something long. Then we see it's a white man and we realize what he picked up is, in fact, a (the signer's) chisel (from how he used it). The word order reflects clarification after the fact. That is, the spatial index (3a) and the classifier predicate (TAKE-LONG-ITEM) precede the information about who was in that spatial index and what long item that classifier predicate involves. The MNPs come late for dramatic impact.

Russian Sign Language presents a (partial) exception to generalization two. SOV is found with classifier predicates, whereas SVO is found with plain verbs, as we expect. But SVO is also found with agreeing verbs (Kimmelman, 2012). And Volterra et al. (1984) report for Italian Sign Language that in non-reversible sentences, SOV is used only if the V is a classifier or somehow else incorporates the O. However, they also say that SVO, the unmarked order, can occur under the same conditions (but see Cecchetto et al., 2006 for the analysis of Italian Sign Language as SOV).

GENERALIZATION THREE
The most common sentence type has only one new argument, which precedes the V.

The fact that the lone argument tends to precede the V is shared by (young) spoken languages. What's not shared is a particular strategy that sign languages often employ. Essentially, sign languages tend to put the relevant players on the stage one at a time, focusing our attention with a single spotlight on a single player, then moving that spotlight to a second player, and so forth.

We saw that same strategy in gestural strings and in homesign (discussed in section A Modal Account).

Possibly, this is a visual strategy. While the retina can receive much information (our visual environment is typically cluttered), at a given time, only a small amount of that information can be processed. "Subjectively, giving attention to any one target leaves less available for others" (Desimone and Duncan, 1995, p. 193). By introducing only one argument per predicate, we increase the chance that each argument will get attention, enhancing good communication of the event. Nevertheless, signers can convey information simultaneously with multiple articulators (both hands, various parts of the face). So we are not convinced this is a visual strategy.

More likely, it is a manual strategy. The manual articulators move slowly in comparison to the speech articulators, which means it takes time to set things up. So once we have the stage set, there's no need to keep doing something as uneconomical as repeating information everyone already knows.

GENERALIZATION FOUR
When two MNPs occur in a locational expression that forms a single clause, the larger more immobile objects tend to precede smaller more mobile ones, regardless of theta role or grammatical function.

Among others, Volterra et al. (1984, pp. 35, 38) suggest this ordering is a direct result of the visual modality because larger objects are perceptually more important, a suggestion supported by a study on the order of gestures (not signs) in which participants consistently place a gesture for a larger stationary object before a gesture for a smaller moving one (Gershkoff-Stowe and Goldin-Meadow, 2002). On the other hand, in both existential and locational sentences animate objects tend to precede inanimate ones (although see Coerts, 1994 and Kristoffersen, 2003 for complications), and sometimes these two principles conflict, which is the explanation these studies give for freer word order in existential/locational sentences, and which is the reason why we did not offer a separate generalization about word order in existential/locational sentences in particular.

A sign utterance that conveys relative spatial information about two objects creates that information spatially and, thus, evokes a cognitive representation of those objects in those spatial positions. It appears that, with respect to that evoked representation, sign languages are sensitive to the relevant visual principles. Studies show that perception of small objects (under 10 cm) differs from perception of large objects (Pakhomova, 2000). Further we perceive small objects as moving more quickly than large objects even when they are moving at the same rate (Leibowitz, 1965). So the fact that existential/ locational sentences tend to establish the location of large objects before they establish the location of small objects appears to follow from some property of visual perception.

EXTRA COMMENT ON GENERALIZATION FIVE
O is immediately adjacent to V.

Since this generalization holds of most spoken languages (which we expect, given the existence of VP) and of creoles (i.e.,

young languages) as well as sign languages, pressures common to all sensorimotor systems apply here. But Meir (2002) points out that in Israeli Sign Language a V can agree with its O without agreeing with the S, a situation not found in spoken languages. This suggests that the visual modality adds pressure for a visual unity or coherence of the V and O in sign languages.

CONCLUSIONS ABOUT VISUAL AND MANUAL PRESSURES

Sign languages are subject to the universal pressures on all languages. Some of those pressures are common to auditory and visual sensorimotor systems and, thus, we suggest they motivate parts of universal grammar. But sign languages are also subject to visual and perhaps manual pressures that set them apart from spoken languages. That sign languages should fall together typologically with respect to various aspects of grammar is not a new claim. For example, all sign languages use simultaneous expressions, a fact most often accounted for by the slowness of the manual articulators (Hohenberger, 2007). By recognizing visual pressure on sign order, we can see that sign languages exploit simultaneity not simply because they can (given that spoken languages can, too—Napoli and Sutton-Spence, 2010), nor totally because of the timing needs due to slow articulators, but because by exploiting it they can better align syntax and semantics with a visual coherence that is at the core of signing itself.

Our study argues that all sign languages will organize signs at the sentence level in a similar way partly because that's how all languages would do it, all else being equal, and partly because the visual modality entails creating pictures. Certainly these pictures are iconic in only the most abstract of ways and that iconicity is concentrated in the productive much more than in the frozen part of the lexicon (Klima and Bellugi, 1979; McDonald, 1985; Brennan, 1990; Taub, 2001; Russo, 2005; Cuxac and Sallandre, 2007; Sallandre, 2007; Konrad, 2011), otherwise any sighted person would be able to understand any sign language. Indeed, in the frozen lexicon, many signs are opaque in that their meanings are not guessable at all. And with respect to the others, signs whose meanings "are most directly interpreted from visibly present referents" or "can be shown by pantomimic expression" are more likely to be understood relatively accurately by people who do not know the given sign language than are signs whose meanings involve some kind of "metonymic association" or are "more culturally specific" (Boyes Braem et al., 2002, p. 187).

But once particular frozen lexical items are understood, and once one understands the nature of all the various types of predicates in sign languages, the organization of frozen and productive signs in the visual space and time of a sign sentence can be seen as largely iconic, where recognizing this iconicity calls for analogy, metaphor, metonymy, and other complex cognitive activities (Napoli and Sutton-Spence, 2011). So the signed creation of pictures demands a visual coherency in order to be interpretable, and this demand for visual coherency should be equally high in any sign language.

MANUAL FACTORS

A few of the studies we cite claim that manual considerations are relevant to word order. Nadeau and Desouvrey (1994, p. 156),

in their study of Quebec Sign Language, suggest that SVO is favored for "mechanical" reasons, claiming that any other order would require additional transitional movements between the signs. Fischer (1975) mentions manual reasons for expecting the SOV order of ASL to change to OSV over time. Two studies point out that the O referred to in a handling classifier must immediately precede the classifier predicate (Jantunen, 2008 for Finnish Sign Language, Sze, 2003 for Hong Kong Sign Language in non-reversible sentences). We leave these remarks for future investigation.

IMPLICATIONS

Universal pressures and visual pressures conspire to bring about the generalizations we have found. We promoted the position here that those universal pressures follow from shared characteristics of the auditory and visual sensorimotor systems and we suggested that those shared characteristics are part of the motivation for universal grammar. Further, as visual pressures, in particular, play a stronger role in sign languages than in spoken languages, they mediate the emergence of the grammars of sign languages in such a way that sign languages tend to converge on a shared design that is, in the respects discussed in section Order and the Visual Modality, different from spoken languages.

We conclude that SOV and SVO should be the prevalent orders found in all declarative sentences in sign languages and that V-initial sentences should be restricted to presentational or existence sentences. In all of this, recall that we are talking only about the distribution of MNP arguments with respect to V. In fact, plain verbs are the only type that should show variation among languages in unmarked word order, specifically between SOV and SVO. That's because plain verbs are the only verbs whose phonological shape is not affected in an iconic way by their arguments. And, as it turns out, SOV and/or SVO are the unmarked orders for plain verbs across all the languages in the studies we examined (see remarks at the end of section Generalization Four under Order and the Visual Modality).

The account of sign order in sign languages that arises from our survey of the data in many studies needs to be tested through examination of a large video corpus, something that has not been possible for most linguists thus far. Fortunately, three major data corpora have recently been made available, for British Sign Language (BSL corpus project, discussed in Schembri, 2008), Auslan (Johnston and Schembri, 2007b; Johnston, 2008, 2010), and Sign Language of the Netherlands (Crasborn and Zwitserlood, 2008). Similar databases are under construction, including for German Sign Language (Hanke et al., 2010), Italian Sign Language (Branchini et al., 2009), Chinese Sign Language (Zhang et al., 2013), and French Belgian Sign Language (Meurant and Sinte, 2013). These databases can serve as a model for building databases for other sign languages, and they will enable researchers to make headway on linguistic analysis with confidence in the foundation upon which arguments are constructed and to both pose and answer questions regrettably infeasible without such a base. We offer our remarks here then, as a starting point for examining sign order with the goal of understanding better the sensorimotor system pressures affecting that order.

AUTHOR CONTRIBUTIONS

ACKNOWLEDGMENTS

REFERENCES

Aboh, E., Pfau, R., and Zeshan, U. (2005). "When a Wh-word is not a Wh-word: the case of Indian Sign Language," in *The Yearbook of South Asian Languages and Linguistics*, ed R. Singh (Berlin: Walter de Gruyter), 11–43.

Akach, P. A. O. (1992). "Sentence formation in Kenyan Sign Language," in *The East African Sign Language Seminar*, eds P. Akach, G. Mweri, and L. Hempel (Copenhagen: Danish Deaf Association), 45–51.

Alexiadou, A., and Anagnostopoulou, E. (1998). Parametrizing AGR: word order, V-movement and EPP-checking. *Nat. Lang. Linguist. Theory* 16, 491–539. doi: 10.1023/A:1006090432389

Antinucci, F., Duranti, A., and Gebert, L. (1979). Relative clause structure, relative clause perception, and the change from SOV to SVO. *Cognition* 7, 145–176. doi: 10.1016/0010-0277(79)90018-0

Aoun, J., Benmamoun, E., and Sportiche, D. (1994). Agreement, word order, and conjunction in some varieties of Arabic. *Linguist. Inq.* 25, 195–220.

Aronoff, M., Meir, I., Padden, C., and Sandler, W. (2003). "Classifier constructions and morphology in two sign languages," in *Perspectives on Classifier Constructions in Sign Languages*, ed K. Emmorey (Mahwah, NJ: Lawrence Erlbaum Associates), 53–84.

Bahan, B. (1996). *Non-Manual Realization of Agreement in American Sign Language*. Boston, MA: Boston University Ph.D. dissertation. Available online at: http://www.bu.edu/asllrp/dissertatns.long.html (Accessed October 26, 2013).

Baker, C., and Cokely, D. (1980). *American Sign Language: A Teacher's Resource Text on Grammar*. Silver Spring, MD: T.J. Publishers.

Baker, M. C. (2012). On the relationship of object agreement and accusative case: evidence from Amharic. *Linguist. Inq.* 43, 255–274. doi: 10.1162/LING_a_00085

Baron, I. (1998). Objet effectué et constructions nominales. *Trav. Linguist.* 35, 115–125.

Bergman, B., and Wallin, L. (1985). "Sentence structure in Swedish Sign Language," in *Sign Language Research '83*, eds W. Stokoe and V. Volterra (Silver Spring, MD: Linstok Press), 217–225.

Bickerton, D. (1981). *Roots of Language*. Ann Arbor, MI: Karoma.

Bickerton, D. (1990). *Language and Species*. Chicago: University of Chicago Press.

Bobaljik, J. D., and Carnie, A. H. (1992). "A Minimalist approach to some problems of irish word order," in *Proceedings of the Harvard Celtic Colloquium*, Vol. 12, eds B. Hillers and J. Hunter (Cambridge, MA: Department of Celtic Languages and Literatures, Harvard University), 110–134.

Bobillo-García, N., Morales-López, E., Reigosa-Varela, C., Freire-Rodríguez, C., and Pérez-Casanova, C. (2006). *El Orden de los Constituyentes de la Oración Declarativa en la LSE*. VI Congreso de Lingüística General, Universidad de Santiago de Compostela. Available online at: http://www.ruc.udc.es/dspace/bitstream/2183/644/1/OPCongrSantiago04.doc. (Accessed October 26, 2013).

Bogaerde, E. M. van den, and Mills, A. E. (1994). "Word order in language input to children: SLN or Dutch," in *Word-order Issues in Sign Language*, eds M. Brennan and G. Turner (Durham: International Sign Linguistics Association), 133–157.

Borsley, R. D., and Roberts, I. (eds.). (2005). "Introduction" in *The Syntax of the Celtic Languages: A Comparative Perspective*, eds R. D. Borsley and I. Roberts (Cambridge: Cambridge University Press), 1–52.

Bouchard, D. (1996). Sign language and language universals: the status of order and position in grammar. *Sign Lang. Stud.* 91, 101–160. doi: 10.1353/sls.1996.0018

Bouchard, D., and Dubuisson, C. (1995). Grammar, order and the position of wh-signs in Quebec Sign Language. *Sign Lang. Stud.* 87, 99–139. doi: 10.1353/sls.1995.0007

Boyes Braem, P., Fournier, M.-L., Rickli, F., Corrazza, S., Franchi, M.-L., and Volterra, V. (1990). "A comparison of techniques for expressing semantic roles and locative relations in two different sign languages," in *SLR '87: Papers from the Fourth International Symposium on Sign Language Research*, eds W. Edmondson and F. Karlsson (Hamburg: Signum-Verlag), 114–120.

Boyes Braem, P., Pizzuto, E., and Volterra, V. (2002). "The interpretation of signs by (hearing and deaf) members of different cultures," in *Progress in Sign Language Research/Fortschritte in der Gebärdensprachforschung*, eds R. Schulmeister and H. Reinitzer (Hamburg: Signum Verlag), 187–220.

Branchini, C., Geraci, C., Giudice, S., and Mereghetti, E. (2009). "Linguistic issues in building a corpus for LIS (Italian Sign Language)," in *Sign Language Corpora: Linguistic Issues Workshop 2009 in London, British Sign Language Corpus Project*. Available online at: http://www.bslcorpusproject.org/wp-content/uploads/slides-sl-corpora-london.pdf (Accessed October 26, 2013).

Braze, D. (2004). Aspectual inflection, verb raising, and object fronting in American Sign Language. *Lingua* 114, 29–58. doi: 10.1016/S0024-3841(03)00034-2

Brennan, M. (1990). *Word Formation in British Sign Language*. Stockholm: Stockholm University Press.

Brennan, M. (1994). "Word order: introducing the issues," in *Word-order Issues in Sign Language*, eds M. Brennan and G. Turner (Durham: International Sign Linguistics Association), 9–45.

Brew, C., and Schulte im Walde, S. (2002). "Spectral clustering for German verbs," in *Proceedings of the Conference on Empirical Methods in Natural Language Processing '02*, eds J. Hagiè and Y. Matsumoto (Stroudsburg, PA: Association for Computational Linguistics), 117–124.

Canale, M. (1978). *Word Order Change in OE: Base Reanalysis in Generative Grammar*. Doctoral dissertation, University of Toronto.

Carnie, A., and Guilfoyle, E. (eds.). (2000). *The Syntax of Verb Initial Languages*. Oxford: Oxford University Press.

Carston, R. (2002). *Thoughts and Utterances: The Pragmatics of Explicit Communication*. Oxford: Blackwell Publishers. doi: 10.1002/9780470754603

Cecchetto, C., Geraci, C., and Zucchi, S. (2006). Strategies of relativization in Italian Sign Language. *Nat. Lang. Linguist. Theory* 24, 945–975. doi: 10.1007/s11049-006-9001-x

Chen Pichler, D. (2011). "Using early ASL word order to shed light on word order variability in sign language," in *Variation in the Input: Studies in The Acquisition of Word Order* (*Studies in Theoretical Psycholinguistics*, Vol. 39), eds M. Anderssen, K. Bentzen, and M. Westergaard (The Netherlands: Springer), 157–177. doi: 10.1007/978-90-481-9207-6_7

Chen Pichler, D., Schalber, K., Hochgesang, J., Milkovic, M., Wilbur, R., Vulje, M., et al. (2008). "Possession and existence in three sign languages," in *Sign Languages: spinning and Unraveling the Past, Present and Future, TISLR9, Florianopolis, Brazil, December 2006*, ed R. Müller de Quadros (Petrópolis; Rio de Janeiro: Editora Arara Azul), 440–458.

Choe, H. S. (1987). "An SVO analysis of VSO languages and parameterization: a study of berber," in *Studies in Berber Syntax. Lexicon Project Working Papers #14*, eds M. Guerssel and K. Hale (Cambridge, MA: Center for Cognitive Science, MIT), 121–158.

Chomsky, N. (2013). Problems of projection. *Lingua* 130, 33–49. doi: 10.1016/j.lingua.2012.12.003

Christensen, P., and Tylén, K. (2013). Gestural representation of event structure in dyadic interaction. *Strings* 1140, 1241–1244. Available online at: http://www.academia.edu/3613166/Gestural_representation_of_event_structure_in_dyadic_interaction

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behav. Brain Sci.* 36, 181–204. doi: 10.1017/S0140525X12000477

Clark, E. V. (1978). "Locationals: existential, locative and possessive constructions," in *Universals of Human Language,* Vol. 4 *Syntax,* ed J. E. Greenberg (Stanford, CA: Stanford University Press), 85–126.

Coerts, J. (1994). "Constituent order in Sign Language of the Netherlands," in *Word-order Issues in Sign Language,* eds M. Brennan and G. Turner (Durham: International Sign Linguistics Association), 47–70.

Corrazza, S., Radutzky, E., Santarelli, B., Verdirosi, M. L., Volterra, V., and Zingarini, A. (1985). "Italian Sign Language: general summary of research," in *SLR'83, Sign Language Research,* eds W. Stokoe and V. Volterra (Silver Spring, MD: Linstok Press), 289–298.

Craig, C. G. (1977). *The Structure of Jacaltec.* Austin, TX: University of Texas Press.

Crasborn, O. (2007). How to recognize a clause when you see one. *Sign. Lang. Linguist.* 10, 103–111. doi: 10.1075/sll.10.2.03cra

Crasborn, O., van der Kooija, E., Waters, D., Woll, B., and Mesch, J. (2008). Frequency distribution and spreading behavior of different types of mouth actions in three sign languages. *Sign Lang. Linguist.* 11, 45–67. doi: 10.1075/sll.11.1.04cra

Crasborn, O., and Zwitserlood, I. (2008). "The Corpus NGT: an online corpus for professionals and laymen," in *Construction and Exploitation of Sign Language Corpora. 3rd Workshop on the Representation and Processing of Sign Languages,* eds O. Crasborn, T. Hanke, E. Efthimiou, I. Zwitserlood, and E. Thoutenhoofd (Parijs: ELDA), 44–49.

Cowie, F. (2008). "Innateness and language," in *Stanford Encyclopedia of Philosophy.* Available online at: http://www.science.uva.nl/~seop/archives/sum2010/entries/innateness-language (Accessed April 4, 2014).

Crowley, T. (1992). *An Introduction to Historical Linguistics.* Oxford: Oxford University Press.

Curtiss, S. (1977). *Genie: A Psycholinguistic Study of a Modern-Day "Wild Child."* Boston, MA: Academic Press.

Curtiss, S., Krashen, S. D., Fromkin, V. A., Rigler, D., and Rigler, M. (1973). "Language acquisition after the critical period: Genie as of April, 1973," in *Papers from the Ninth Regional Meeting Chicago Linguistic Society,* ed R. I. Binnick (Chicago: University of Chicago, Department of Linguistics), 98–103.

Cuxac, C. (2000). *La Langue des Signes Française (LSF): les voies de l'iconicité.* Paris: Ophrys.

Cuxac, C., and Sallandre, M-A. (2007). "Iconicity and arbitrariness in French Sign Language: highly iconic structures, degenerated iconicity and grammatic iconicity," in *Verbal and Signed languages: Comparing Structure, Constructs and Methodologies,* eds E. Pizzuto, P. Pietrandrea, and R. Simone (Berlin: de Gruyter), 13–33.

DeGraff, M. (2003). Against creole exceptionalism. *Language* 79, 391–410. doi: 10.1353/lan.2003.0114

DeGraff, M. (2005). "Morphology and word order in 'creolization'and beyond," in *The Oxford Handbook of Comparative Syntax,* eds G. Cinque and R. S. Kayne (Oxford: Oxford University Press), 293–372.

De Langhe, O., Guitteny, P., Portine, H., and Retoré, C. (2004). A propos des structures OSV en langue des signes française. *Silexicales* 4, 115–130. Available online at: http://www.labri.fr/perso/retore/ARTICLES/OSV.rtf.pdf

Delgado-Martins, M. R., Coutinho, A., and Amaral, M. A. (1994). *Para uma Gramática da Língua Gestual Portuguesa.* Lisbon: Editorial Caminho.

Delsing, L.-O. (2000). "From OV to VO in Swedish," in *Diachronic Syntax: Models and Mechanisms,* eds S. Pintzuk, G. Tsoulas, and A. Warner (Oxford: Oxford University Press), 255–274.

Desimone, R., and Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annu. Rev. Neurosci.* 18, 193–222. doi: 10.1146/annurev.ne.18.030195.001205

de Swart, H. (2009). "Negation in early L2: a window on language genesis," in *Language Evolution: The View from Restricted Linguistic Systems,* eds R. Botha and H. de Swart (Utrecht: LOT Publications), 59–100.

Deuchar, M. (1983). "Is BSL an SOV language?" in *Language in Sign: An International Perspective on Sign Language,* eds J. Kyle and B. Woll (London: Croom Helm), 59–76.

Dryer, M. S. (2003) "Word order in Sino-Tibetan languages from a typological and geographical perspective," in *The Sino-Tibetan Languages,* eds G. Thurgood and R. J. LaPolla (London: Routledge), 43–55.

Dryer, M. S. (2005). "Order of subject, object and verb," in *The World Atlas of Language Structures,* eds M. Haspelmath, M. S. Dryer, D. Gil, and B. Comrie (Oxford: Oxford University Press), 330–333.

Efthimiou, E., and Fotinea, S-E. (2007). "GSLC: creation and annotation of a Greek Sign Language corpus for HCI," in *Proceedings of the 4th International Conference on Universal Access in Human Computer Interaction: Coping with diversity,* ed C. Stephanidis (Berlin: Springer Verlag), 657–666.

Emmorey, K. (ed.). (2003). *Perspectives on Classifier Constructions in Sign Languages.* Mahwah, NJ: Lawrence Erlbaum Associates.

Engberg-Pedersen, E. (1994). "Some simultaneous constructions in Danish Sign Language," in *Word-Order Issues in Sign Language,* eds M. Brennan and G. Turner (Durham: International Sign Linguistics Association), 73–88.

Farris, M. (1994). Sign language research and Polish Sign Language. *Lingua Posnaniensis* 36, 13–36.

Fedorenko, E., Duncan, J., and Kanwisher, N. (2012). Language-selective and domain-general regions lie side by side within Broca's area. *Curr. Biol.* 22, 2059–2062. doi: 10.1016/j.cub.2012.09.011

Fischer, S. (2010). Word-order Change as a Source of Grammaticalisation. Amsterdam: John Benjamins. doi: 10.1075/la.157

Fischer, S. D. (1975). "Influences on word order change in American Sign Language," in *Word Order and Word Order Change,* ed C. Li (Austin, TX: University of Texas Press), 3–25.

Fischer, S. D. (1990). "The Head Parameter in ASL," in *SLR'87: Papers from the Fourth International Symposium on Sign Language Research,* eds W. Edmondson and F. Karlsson (Hamburg: Signum Press), 75–85.

Fischer, S. D., and Janis, W. (1990). "Verb sandwiches in ASL," in *Current Trends in European Sign Language Research: Proceedings of the 3rd European Congress on Sign Language Research, Hamburg, 1989,* eds S. Prillwitz and T. Vollhaber (Hamburg: Signum Press), 279–294.

Flack Potts, K. (2007). "Ambiguity avoidance as contrast preservation: case and word order freezing in Japanese," in *UMass Occasional Papers in Linguistics 32: Papers in Optimality Theory III,* eds L. Bateman, M. O'Keefe, E. Reilly, and A. Werle (Amherst, MA: BookSurge Publishing), 57–89.

Friedman, L. (1976). "The manifestation of subject, object, and topic in American Sign Language," in *Subject and Topic,* ed C. Li (New York, NY: Academic Press), 125–148.

Fromkin, V., Krashen, S., Curtiss, S., Rigler, D., and Rigler, M. (1974). The development of language in Genie: A case of language acquisition beyond the "critical period". *Brain Lang.* 1, 81–107. doi: 10.1016/0093-934X(74)90027-3

Gallese, V., Gernsbacher, M. A., Heyes, C., Hickok, G., and Iacoboni, M. (2011). Mirror neuron forum. *Perspect. Psychol. Sci.* 6, 369–407. doi: 10.1177/1745691611413392

Gell-Mann, M., and Ruhlen, M. (2011). The origin and evolution of word order. *Proc. Natl. Acad. Sci. U.S.A.* 108, 17290–17295. doi: 10.1073/pnas.1113716108

Gershkoff-Stowe, L., and Goldin-Meadow, S. (2002). Is there a natural order for expressing semantic relations? *Cogn. Psychol.* 45, 375–412. doi: 10.1016/S0010-0285(02)00502-9

Gibson, E., and Fedorenko, E. (2013). The need for quantitative methods in syntax and semantics research. *Lang. Cogn. Process.* 28, 88–124. doi: 10.1080/01690965.2010.515080

Gibson, E., Piantadosi, S. T., Brink, K., Bergen, L., Lim, E., and Saxe, R. (2013). A noisy-channel account of crosslinguistic word order variation. *Psychol. Sci.* 24, 1079–1088. doi: 10.1177/0956797612463705

Giles, J. (2006). The trouble with replication. *Nat. Lang.* 442, 344–347. doi: 10.1038/442344a

Givón, T. (1975). "Serial verbs and syntactic change: Niger-Congo," in *Word Order and Word Order Change,* ed C. Li (Austin, TX: University of Texas Press), 47–112.

Givón, T. (1976). "Topic, pronoun, and grammatical agreement," in *Subject and Topic,* ed C. N. Li (New York, NY: Academic Press), 149–188.

Givón, T. (1979). *On Understanding Grammar.* New York, NY: Academic Press.

Glück, S., and Pfau, R. (1998). "On classifying classification as a class of inflection in German Sign Language," in *Proceedings of ConSole 6,* eds T. Cambier-Langeveld, A. Lipták, and M. Redford (Leiden: SOLE), 59–74.

Goldberg, A. (1995). *Constructions: A Construction Grammar Approach to Argument Structure.* Chicago: University of Chicago Press.

Goldberg, A. (2003). Constructions: a new theoretical approach to language. *Trends Cogn. Sci.* 7, 219–224. doi: 10.1016/S1364-6613(03)00080-9

Goldin-Meadow, S. (1978). A study in human capacities. *Science* 200, 649–651. doi: 10.1126/science.200.4342.649

Goldin-Meadow, S. (2003). *The Resilience of Language: What Gesture Creation in Deaf Children Can Tell us About How all Children Learn Language*. New York, NY: Psychology Press.

Goldin-Meadow, S., So, W. C., Özyürek, A., and Mylander, C. (2008). The natural order of events: how speakers of different languages represent events nonverbally. *Proc. Natl. Acad. Sci. U.S.A.* 105, 9163–9168. doi: 10.1073/pnas.0710060105

Hale, K. (1975). "Gaps in Grammar and Culture," in *Linguistics and Anthropology: In Honor of C.F. Voegelin*, eds M. D. Kinkade, K. L. Hale, and O. Werner (Lisse: The Peter de Ridder Press), 295–315.

Hall, M. L., Mayberry, R. I., and Ferreira, V. S. (2013). Cognitive constraints on constituent order: evidence from elicited pantomime. *Cognition* 129, 1–17. doi: 10.1016/j.cognition.2013.05.004

Hänel, B. (2005). The acquisition of agreement in DGS: early steps into spatially expressed syntax. *Linguist. Ber.* 13, 201–232.

Hanke, T., König, L., Wagner, S., and Matthes, S. (2010). "DGS corpus and dicta-sign: the Hamburg studio setup," in *7th International Conference on Language Resources and Evaluation. Workshop Proceedings. W13, 4th Workshop on the Representation and Processing of Sign Languages: Corpora and Sign Language Technologies*, ed P. Dreuw (Paris: ELRA), 106–109.

Happ, D., and Verköper, M. O. (2006). *Deutsche Gebärdensprache: Ein Lehr-und arbeitsbuch*. Frankfurt: Fachhochschulverlag.

Haviland, J. B. (2011). "Nouns, verbs, and constituents in an emerging 'Tzotzil' sign language," in *Representing Language: Essays in Honor of Judith Aissen*, eds R. Gutiérrez-Bravo, L. Mikkelsen, and E. Potsdam (Santa Cruz, CA: Linguistics Research Center), 157–171.

Haywood, S., Pickering, M., and Branigan, H. (2005). Do speakers avoid ambiguity during dialogue? *Psychol. Sci.* 16, 362–366. doi: 10.1111/j.0956-7976.2005.01541.x

Herring, S. C., and Paolillo, J. C. (1995). "Focus position in SOV languages," in *Word Order in Discourse*, eds P. A. Downing and M. Noonan (Amsterdam: John Benjamins), 163–198. doi: 10.1075/tsl.30.08her

Hoekstra, T., and Mulder, R. (1990). Unergatives as copular verbs; locational and existential predication. *Linguist. Rev.* 7, 1–80. doi: 10.1515/tlir.1990.7.1.1

Hohenberger, A. (2007). "The possible range of variation between sign languages: Universal Grammar, modality, and typological aspects," in *Visible Variation*, eds W. Bisang, H. H. Hock, and W. Winter (Berlin; New York: Mouton de Gruyter), 341–383.

Hróarsdóttir, Þ. (2000). *Word Order Change in Icelandic: From OV to VO*. Amsterdam: John Benjamins.

Jantunen, T. (2008). Fixed and free: order of the verbal predicate and its core arguments in declarative transitive clauses in Finnish Sign Language. *SKY J. Linguist.* 21, 83–123. Available online at: http://www.linguistics.fi/julkaisut/SKY2008/Jantunen_NETTIVERSIO.pdf

Jefwa, G. J. (2009). Structural borrowing: the case of Kenyan Sign Language (KSL) and Kiswahili contact signing. *J. Lang. Technol. Entrep. Afr.* 1, 160–174. doi: 10.4314/jolte.v1i2.41781

Johnson, J. S., and Newport, E. (1989). Critical period effects in second language learning: the influence of maturational state on the acquisition of English as a second language. *Cogn. Psychol.* 21, 60–99. doi: 10.1016/0010-0285(89)90003-0

Johnston, T. (2008). *Corpus of Grammar and Discourse Strategies of Deaf Native Users of Auslan (Australian Sign Language), Endangered Languages Archive*. London: SOAS, University of London. Available online at: http://elar.soas.ac.uk/node/3 (Accessed 2012).

Johnston, T. (2010). From archive to corpus: transcription and annotation in the creation of signed language corpora. *Int. J. Corpus Linguist.* 15, 106–131. doi: 10.1075/ijcl.15.1.05joh

Johnston, T. (2013). Formational and functional characteristics of pointing signs in a corpus of Auslan (Australian sign language): are the data sufficient to posit a grammatical class of 'pronous' in Auslan? *Corpus Linguist. Linguist. Theory* 9, 109–159. doi: 10.1515/cllt-2013-0012

Johnston, T., and Schembri, A. (2007a). *Australian Sign Language (Auslan): An Introduction to Sign Language Linguistics*. Cambridge: Cambridge University Press.

Johnston, T., and Schembri, A. (2007b). "Testing language description through language documentation, archiving and corpus creation: the case of indicating verbs in the Auslan Archive Corpus," in *Proceedings of Conference on Language Documentation and Linguistic Theory*, eds P. K. Austin, O. Bond, and D. Nathan (London: SOAS, University of London), 145–154.

Johnston, T., Vermeerbergen, M., Schembri, A., and Leeson, L. (2007). "'Real data are messy': considering cross-linguistic analysis of constituent ordering in Auslan, VGT, and ISL," in *Visible Variation: Comparative Studies on Sign Language Structure*, eds P. Perniss, R. Pfau, and M. Steinbach (Berlin: Mouton), 163–206.

Julien, M. (2002). *Syntactic Heads and Word Formation*. Oxford: Oxford University Press.

Kegl, J. (1990). "Predicate argument structure and verb-class organization in the ASL lexicon," in *Sign Language Research: Theoretical Issues*, ed C. Lucas (Washington, DC: Gallaudet University Press), 149–175.

Kemmerer, D. (2012). The cross-linguistic prevalence of SOV and SVO word orders reflects the sequential and hierarchical representation of action in Broca's Area. *Lang. Linguist. Compass* 6, 50–66. doi: 10.1002/lnc3.322

Khanittanan, W. (a.k.a. Kanittanan) (1986). "Kamti Tai: from an SVO to an SOV language," *South Asian Languages: Structure, Convergence and Diglossia*, ed B. H. Krishnamurti (Delhi: Motilal Banarsidass), 174–178.

Kimmelman, V. (2012). Word order in Russian Sign Language. *Sign Lang. Stud.* 12, 414–445. doi: 10.1353/sls.2012.0001

Klima, E., and Bellugi, U. (1979). *The Signs of Language*. Cambridge, MA: Harvard University Press.

Konrad, R. (2011). Die lexikalische Struktur der Deutschen Gebärdensprache im Spiegel empirischer Fachgebärdenlexikographie. Tübingen: Narr Verlag

Koutsoudas, A. (1969). *Workbook in Syntax*. New York, NY: McGraw Hill.

Kristoffersen, J. H. (2003). "Existence, location and possession and the order of constituents in Danish Sign Language," in *Cross-linguistic Perspectives in Sign language Research: Selected Papers from TISLF 2000*, eds A. Baker, B. E. M. van den Bogaerde, and O. Crasborn (Hamburg: Signum), 131–139.

Kuhn, J. (2001). *Formal and Computational Aspects of Optimality-Theoretic Syntax*. Stuttgart: Universität Stuttgart, Institut für maschinelle Sprachverarbeitung.

Langus, A., and Nespor, M. (2010). Cognitive systems struggling for word order. *Cogn. Psychol.* 60, 291–318. doi: 10.1016/j.cogpsych.2010.01.004

Lee, H. (2001). "Markedness and word order freezing," in *Formal and Empirical Issues in Optimality Theoretic Syntax*, ed P. Sells (Stanford, CA: CSLI), 63–128.

Leeson, L., and Saeed, J. (2012). "Word order," in *Sign Language. An International Handbook (HSK—Handbooks of Linguistics and Communication Science)*, eds R. Pfau, M. Steinbach, and B. Woll (Berlin: Mouton de Gruyter), 245–264.

Leibowitz, H. (1965). *Visual Perception*. New York, NY: Macmillan.

Leuninger, H. (2000). "Mit den Augen lernen: Gebärdenspracherwerb," in *Enzyklopädie der Psychologie,* Vol. 3. *Sprachentwicklung*, ed H. Grimm (Göttingen: Hogrefe), 229–270.

Levy, R. (2008). "A noisy-channel model of rational human sentence comprehension under uncertain input," in *Proceedings of the Conference on Empirical Methods in Natural Language Processing* (Stroudsburg, PA: Association for Computational Linguistics), 234–243. doi: 10.3115/1613715.1613749

Levy, R., and Jaeger, T. F. (2007). "Speakers optimize information density through syntactic reduction," in *Advances in Neural Information Processing Systems 19*, eds B. Schölkopf, J. Platt, and T. Hoffman (Cambridge, MA: MIT Press), 849–856.

Liddell, S. (1990). "Four functions of a locus: reexamining the structure of space in ASL," in *Sign Language Research: Theoretical Issues*, ed C. Lucas (Washington, DC: Gallaudet University Press), 176–198.

Liddell, S., and Johnson, R. (1986). American Sign Language compound formation processes, lexicalization, and phonological remnants. *Nat. Lang. Linguist. Theory* 4, 445–513. doi: 10.1007/BF00134470

Lillo-Martin, D. (1986). Two kinds of null arguments in American Sign Language. *Nat. Lang. Linguist. Theory* 4, 415–444. doi: 10.1007/BF00134469

Lillo-Martin, D. (1991). *Universal Grammar and American Sign Language: Setting the Null Argument Parameters*. Dordrecht: Kluwer.

Lillo-Martin, D., and Berk, S. (2003). "Acquisition of constituent order under delayed language exposure," in *Proceedings of the 27th Annual Boston University Conference on Language Development*, Vol. 2 (Somerville, MA: Cascadilla Press), 484–495.

Massone, M. I., and Curiel, M. (2004). Sign order in Argentine Sign Language. *Sign Lang. Stud.* 5, 63–93. doi: 10.1353/sls.2004.0023

Matsuoka, K. (1997). Verb raising in American Sign Language. *Lingua* 103, 127–149. doi: 10.1016/S0024-3841(97)00015-6

McDonald, B. (1985). "Productive *on Sign* and frozen lexicon in ASL: an old problem revisited," in *SLR'83: Proceedings of the Third International Symposium Language Research, Rome, June 22–26, 1983*, eds W. Stokoe and V. Volterra (Silver Spring, MD: Linstok), 254–259.

McEnery, T., and Wilson, A. (1996). *Corpus Linguistics.* Edinburgh: Edinburgh University Press.

McIntire, M. (1980). *Locatives in American Sign Language.* Los Angeles, CA: University of California.

McKee, R., and Kennedy, G. (2005). "New Zealand Sign Language," in *Languages of New Zealand*, eds A. Bell, R. Harlow, and D. Starks (Wellington, NZ: Victoria University Press), 271–297.

McWhorter, J. H. (ed.). (2000). *Language Change and Language Contact in Pidgins and Creoles.* Amsterdam: John Benjamins. doi: 10.1075/cll.21

Meier, R., Cormier, K., and Quinto-Pozos, D. (2002). *Modality and Structure in Signed and Spoken Languages.* Cambridge: Cambridge University Press.

Meir, I. (1995). "Explaining backwards verbs in ISL: syntactic-semantic interaction," in *Proceedings of the 4th European Congress on Sign Language Research*, eds H. Bos and T. Schermer (Hamburg: Signum), 105–119.

Meir, I. (2002). A cross-modality perspective on verb agreement. *Nat. Lang. Linguist. Theory* 20, 413–450. doi: 10.1023/A:1015041113514

Meir, I., Lifshitz, A., Ilkbasaran, D., and Padden, C. (2010a). "The interaction of animacy and word order in human languages: a study of strategies in a novel communication task," in *Proceedings of the Eighth Evolution of Language Conference*, eds A. D. M. Smith, B. Schouwstra, B. de Boer, and K. Smith (Singapore: World Scientific Publishing Co.), 455–456

Meir, I., Padden, C., Aronoff, M., and Sandler, W. (2007). Body as subject. *J. Linguist.* 43, 531–563. doi: 10.1017/S0022226707004768

Meir, I., and Sandler, W. (2007). *A Language in Space: The story of Israeli Sign Language.* New York, NY: Lawrence Erlbaum.

Meir, I., Sandler, W., Padden, P., and Aronoff, M. (2010b) "Emerging sign languages," in *Oxford Handbook of Deaf Studies, Language, and Education*, Vol. 2, eds M. Marschark and P. E. Spencer (Oxford: Oxford University Press), 267–280.

Meurant, L., and Sinte, A. (2013). "Toward a corpus of French Belgian Sign Language (LSFB) discourses," in *Across the Line of Speech and Writing Variation, Corpora and Language in Use –Proceedings 2,* eds C. Bolly and L. Degand (Louvain-la-Neuve: Presses Universitaires de Louvain), 199–212.

Miljan, M. (2000). *The Noun Phrase in Estonian Sign Language from the Typological Perspective.* Tallinn: Estonian Institute of Humanities.

Milkovic, M., Bradaric-Joncic, S., and Wilbur, R. (2006). Word order in Croatian Sign Language. *Sign Lang. Linguist.* 9, 169–206. doi: 10.1075/sll.9.1-2.10mil

Milkovic, M., Bradaric-Joncic, S., and Wilbur, R. (2007). Information status and word order in Croatian Sign Language. *Clin. Linguist. Phon.* 21, 1007–1017. doi: 10.1080/02699200701659227

Minoura, N. (2008). Word order in Malagasy Sign Language. 東京外国語大学論集(*Journal of Tokyo University of Foreign Studies*) 77, 47–69. Available online at: http://repository.tufs.ac.jp/bitstream/10108/70849/2/acs084009_ful.pdf

Morales-López, E., Boldú-Menasanch, R. M., Alonso-Rodríguez, J. A., Gras-Ferrer, V., and Rodríguez-González, M. A. (2005). The verbal system of Catalan Sign Language. *Sign Lang. Stud.* 5, 441–496. doi: 10.1353/sls.2005.0018

Morales-López, E., Reigosa-Varela, C., and Bobillo-García, N. (2012). Word order and informative functions (topic and focus) in Spanish Signed Language utterances. *J. Pragmat.* 44, 474–489. doi: 10.1016/j.pragma.2011.12.010

Moravcsik, E. (1988). "Agreement and markedness," in *Agreement in Natural Language: Approaches, Theories, Descriptions*, eds M. Barlow and C. A. Ferguson (Stanford, CA: CSLI), 89–106.

Nadeau, M., and Desouvrey, L. (1994). "Word order in sentences with directional verbs in Quebec Sign Language," in *Perspectives on Sign Language Structure. Papers from the Fifth International Symposium on Sign Language Research*, eds I. Ahlgren, B. Bergman, and M. Brennan (Durham: International Sign Linguistics Association), 149–158.

Nakanashi, K. (1994). "The influence of Japanese word order on Japanese Sign Language," in *Word-order Issues in Sign Language*, eds M. Brennan and G. Turner (Durham: International Sign Linguistics Association), 171–192.

Napoli, D. J., and Sutton-Spence, R. (2010). Limitations on simultaneity in sign language. *Language* 86, 647–662. doi: 10.1353/lan.2010.0018

Napoli, D. J., and Sutton-Spence, R. (2011). "Sign language humor, human singularities, and the origins of language," in *Deaf Around the World: The Impact of*

Language, eds G. Mathur and D. J. Napoli (Oxford: Oxford University Press), 231–250.

Neidle, C., Kegl, J., MacLaughlin, D., Bahan, B., and Lee, R. (2000). *The Syntax of American Sign Language: Functional Categories and Hierarchical Structure.* Cambridge MA: MIT Press.

Newmeyer, F. (2000a). *Language Form and Language Function.* Cambridge, MA: MIT Press.

Newmeyer, F. (2000b). O"n the reconstruction of 'proto-world' word order," in *The Evolutionary Emergence of Language: Social Function and the Origins of Linguistic Form*, eds C. Knight, M. Studdert-Kennedy, and J. Hurford (Cambridge: Cambridge University Press), 372–388. doi: 10.1017/CBO9780511606441.022

Newport, E. (1990). Maturational constraints on language learning. *Cogn. Sci.* 14, 11–28. doi: 10.1207/s15516709cog1401_2

Newport, E. (1991). "Contrasting concepts of the critical period for language," in *The Epigenesis of Mind: Essays on Biology and Cognition*, eds S. Carey and R. Gelman (New York, NY: Lawrence Erlbaum), 111–132.

Newport, E., Bavelier, D., and Neville, H. J. (2001). "Critical thinking about critical periods: perspectives on a critical period for language acquisition," in *Language, Brain, and Cognitive Development: Essays in Honor of Jacques Mehler*, ed E. Dupoux (Cambridge, MA: MIT Press), 481–502.

Nyst, V. (2007). *A Descriptive Analysis of Adamorobe Sign Language.* Amsterdam: University of Amsterdam.

Ochs, E. (1979). "Transcription as theory," in *Developmental Pragmatics*, eds E. Ochs and B. Schieffelin (New York, NY: Academic Press), 43–72.

Oviedo, A. (2001). *Apuntes Para una Gramática de la Lengua de Señas Colombiana.* Santa Fe de Bogotá: Instituto Nacional para Sordos and the Universidad del Valle.

Padden, C. (1988). *Interaction of Morphology and Syntax in American Sign Language.* New York, NY: Garland Publishing.

Padden, C. (1990). "The relation between space and grammar in ASL verb morphology," in *Sign Language Research: Theoretical Issues*, ed C. Lucas (Washington, DC: Gallaudet University Press), 118–132.

Padden, C., Meir, I., Sandler, W., and Aronoff, M. (2010). "Against all expectations: encoding subjects and objects in a new language," in *Hypothesis A/Hypothesis B: Linguistic Explorations in Honor of David M. Perlmutter*, eds D. Gerdts, J. Moore, and M. Polinsky (Cambridge, MA: MIT Press), 383–400.

Pakhomova, A. S. (2000). Interhemisphere asymmetry and the constancy of visual perception of small and large sizes. *Hum. Physiol.* 26, 275–280. doi: 10.1007/BF02760187

Parodi, C. (1995). "Participle agreement and object shift in Old Spanish: a minimalist theory," in *Evolution and Revolution in Linguistic Theory*, eds H. Campos and P. M. Kempchinsky (Washington, DC: Georgetown University Press), 276–301.

Petitto, L. A., Zatorre, R. J., Gauna, K., Nikelski, E. J., Dostie, D., and Evans, A. C. (2000). Speech-like cerebral activity in profoundly deaf people processing signed languages: implications for the neural basis of human language. *Proc. Natl. Acad. Sci.* 97, 13961–13966. doi: 10.1073/pnas.97.25.13961

Petronio, K. (1993). *Clause Structure in American Sign Language.* Seattle, WA: University of Washington Ph.D. dissertation.

Plaza-Pust, C. (2008). "Why variation matters: on language contact in the development of L2 written German," in *Sign Bilingualism: Language Development, Interaction, and Maintenance in Sign Language Contact Situations*, eds C. Plaza-Pust and E. Morales-López (Amsterdam: John Benjamins), 73–135. doi: 10.1075/sibil.38.07pla

Plaza-Pust, C., and Weinmeister, K. (2008). "Bilingual acquisition of German Sign Language and written German: developmental asynchronies and language contact," in *Sign Languages: Spinning and Unraveling the Past, Present and Future, TISLR9, Florianopolis, Brazil, December 2006*, ed R. M. de Quadros (Petrópolis/Rio de Janeiro: Editora Arara Azul), 497–529.

Quadros, R. M. de. (2003). "Phrase structure of Brazilian Sign Language," in *Cross-linguistic Perspectives in Sign Language Research: Selected Papers from TISLF 2000*, eds A. Baker, B. E. M. van den Bogaerde, and O. Crasborn (Hamburg: Signum), 141–161.

Quadros, R. M. de, and Lillo-Martin, D. (2010). "Clause structure," in *Sign Languages*, ed D. Brentari (Cambridge: Cambridge University Press), 225–251.

Quadros, R. M. de, and Quer, J. (2008). "Back to back(wards) and moving on: on agreement, auxiliaries and verb classes in sign language," in *Sign Languages: Spinning and Unraveling the Past, Present and Future, TISLR9, Florianopolis, Brazil, December 2006*, ed R. M. de Quadros (Petrópolis/Rio de Janeiro: Editora Arara Azul), 530–551.

Quinto, D. (1999). *Word Order in Mexican Sign Language*. Austin, TX: University of Texas at Austin. Available by writing to the author.

Raney, R. (1984). "VSO and SVO Order in Welsh and Breton" in *Papers and Studies in Contrastive Linguistics*, Vol. 18, eds A. K. Torbus and W. Z. Okolice (Poznan: Adam Mickiewicz Univ), 47–54.

Roberge, P. T. (2009). "The creation of pidgins as a possible window on language evolution," in *Language Evolution: The View from Restricted Linguistic Systems*, eds R. Botha and H. de Swart (Utrecht: LOT Publications), 101–138.

Rosenbaum, D. A. (2009). *Human Motor Control*. New York, NY: Academic Press.

Rosenbaum, D. A., Carlson, R. A., and Gilmore, R. O. (2001). Acquisition of intellectual and perceptual-motor skills. *Annu. Rev. Psychol.* 52, 453–470. doi: 10.1146/annurev.psych.52.1.453

Rosenstein, O. (2004). "Topics in ISL (Israeli Sign Language) are not topicalized," in *Poster Presentation at the 8th International Conference on Theoretical Issues in Sign Language Research (TISLR)* (Barcelona). Available online at: http://www.ub.edu/ling/tislr8/Rosenstein.doc (Accessed October 27, 2013).

Ross, M. (2001). "Contact-induced change in Oceanic languages in North-West Melanesia," in *Areal Diffusion and Genetic Inheritance: Problems in Comparative Linguistics*, eds A. Y. Aikhenvald and R. M. W. Dixon (Oxford: Oxford University Press), 134–66.

Russo, T. (2005). A crosslinguistic, cross-cultural analysis of metaphors in two Italian Sign Language (LIS) registers. *Sign Lang. Stud.* 5, 333–359. doi: 10.1353/sls.2005.0009

Sallandre, M. (2007). "Simultaneity in French Sign Language discourse," in *Simultaneity in Signed Languages: Form and Function*, eds M. Vermeerbergen, L. Leeson, O. Crasborn, W. Sandler, and D. Lillo-Martin (Amsterdam: John Benjamins), 103–126.

Sandler, W., Meir, I., Padden, C., and Aronoff, M. (2005). The emergence of grammar: systematic structure in a new language. *Proc. Natl. Acad. Sci. U.S.A.* 102, 2661–2665. doi: 10.1073/pnas.0405448102

Schembri, A. (2008). "British Sign Language corpus project: open access archives and the observer's paradox," in *Sixth International Conference on Language Resources and Evaluation, May 26-June 1, 2008* (Marrakech). Available online at: http://www.lrec-conf.org/proceedings/lrec2008/ (Accessed October 26, 2013).

Schouwstra, M. (2012). *Semantic Structures, Communicative Strategies, and the Emergence of Language*. Utrecht: LOT.

Schuit, J., Baker, A., and Pfau, R. (2011). Inuit Sign Language: a contribution to sign language typology. *Linguist. Amst.* 4, 1–31. Available online at: http://www.linguisticsinamsterdam.nl/04/nr01/a01

Senghas, A., Coppola, M., Newport, E. L., and Supalla, T. (1997). "Argument structure in Nicaraguan Sign Language: the emergence of grammatical devices," in *Proceedings of the Boston University Conference on Language Development*, eds E. Hughes, M. Hughes, and A. Greenhill (Boston, MA: Cascadilla Press), 550–561.

Shannon, C. (1948). A mathematical theory of communication. *Bell Syst. Tech. J.* 27, 623–656. doi: 10.1002/j.1538-7305.1948.tb00917.x

Slobin, D. I. (2006). "Issues of linguistic typology in the study of sign language development of deaf children," in *Advances in the Sign Language Development of Deaf Children*, eds B. Schick, M. Marschark, and P. Spencer (Oxford: Oxford University Press), 20–45.

Smith, W. (2005). Taiwan Sign Language research: an historical overview. *Lang. Linguist.* 6, 187–215. Available online at: http://www.ling.sinica.edu.tw/files/publication/j2005_2_02_5817.pdf

So, W. C., Coppola, M., Licciardello, V., and Goldin-Meadow, S. (2005). The seeds of spatial grammar in the manual modality. *Cogn. Sci.* 29, 1029–1043. doi: 10.1207/s15516709cog0000_38

Sprouse, J. (2011). A test of the cognitive assumptions of magnitude estimation: commutativity does not hold for acceptability judgments. *Language* 87, 274–288. doi: 10.1353/lan.2011.0028

Sundquist, J. D. (2006). Syntactic variation in the history of Norwegian and the decline of XV word order. *Diachronica* 23, 105–141. doi: 10.1075/dia.23.1.06sun

Sutton-Spence, R., and Woll, B. (1999). *The Linguistics of British Sign Language: An Introduction*. Cambridge: Cambridge University Press.

Sze, Y. B. F. (2003). "Word order of Hong Kong Sign Language," in *Cross-linguistic Perspectives in Sign Language Research: Selected Papers from TISLF 2000*, eds A. Baker, B. E. M. van den Bogaerde, and O. Crasborn (Hamburg: Signum), 163–191.

Tallerman, M. (1998). "Word order in Celtic," in *Constituent Order in the Languages of Europe*, ed A. Siewierska (The Hague: Mouton de Gruyter), 21–46.

Taub, S. (2001). *Language from the Body: Iconicity and Metaphor in American Sign Language*. Cambridge: Cambridge University Press. doi: 10.1017/CBO9780511509629

Taylor, A. (1994). The change from SOV to SVO in Ancient Greek. *Lang. Var. Change* 6, 1–37. doi: 10.1017/S0954394500001563

Tettamanti, M., Buccino, G., Saccuman, M. C., Gallese, V., Danna, M., Scifo, P., et al. (2005). Listening to action-related sentences activates fronto-parietal motor circuits. *J. Cogn. Neurosci.* 17, 273–281. doi: 10.1162/0898929053124965

Timm, L. A. (1989). Word order in 20th century Breton. *Nat. Lang. Linguist. Theory* 7, 361–378. doi: 10.1007/BF00208101

Tomlin, R. (1986). *Basic Word Order. Functional Principles*. London: Croom Helm.

Tsay, J., and Myers, J. (2009). "The morphology and phonology of Taiwan Sign Language," in *Taiwan Sign Language and Beyond*, eds J. H.-Y. Tai and J. Tsay (Taipei: Taiwan Institute for the Humanities, National Chung Cheng University), 83–130.

Tomasello, M. (2003). *Constructing a Language: A Usage-Based Theory of Language Acquisition*. Cambridge, MA: Harvard University Press.

Van Gelderen, E. (2011). *The Linguistic Cycle: Language Change and the Language Faculty*. Oxford: Oxford University Press.

Varin, A. (1979). VSO and SVO order in Breton. *Arch. Linguist. Leeds* 10, 83–101.

Vermeerbergen, M. (1996). ROOD KOOL TIEN PERSOON IN. Morfo-syntactische Aspecten van Gebarentaal. [RED CABBAGE TEN PERSON IN. Morpho-syntactic Aspects of Sign Language, Where that Flemish Sign Sentence Means 'That (big) Red Cabbage Can feed 10 People'.] Brussels: Vrije Universiteit Brussel Dissertation.

Vermeerbergen, M., Van Herreweghe, M., and Akach, P. (2007a). "Word order issues in South African Sign Language and Flemish Sign Language," in *Visible Variation: Comparative Studies in Sign Language Structure*, eds P. Perniss, R. Pfau, and M. Steinbach (Berlin: Mouton de Gruyter), 163–205.

Vermeerbergen, M., Van Herreweghe, M., Akach, P., and Matabane, E. (2007b). Constituent order in Flemish Sign Language (VGT) and South African Sign Language (SASL): a cross-linguistic study. *Sign Lang. Linguist.* 10, 25–54. doi: 10.1075/sll.10.1.04ver

Volterra, V., Laudanna, A., Corazza, S., Radutzky, E., and Natale, F. (1984). "Italian Sign Language: the order of elements in the declarative sentence," in *Recent Research on European Sign Language*, eds F. Loncke, P. Boyes Braem, and Y. Lebrun (Lisse: Swets and Zeitlinger BV), 19–46.

Vulanoviæ, R. (2005). The combinatorics of cases and word order. *Res. Lang. Comput.* 3, 107–129. doi: 10.1007/s11168-005-1289-x

Washabaugh, W., Woodward, J., and De Santis, S. (1978). Providence Island Sign Language: a context-dependent language. *Anthropol. Linguist.* 20, 95–109.

Weskott, T., and Fanselow, G. (2011). On the informativity of different measures of linguistic acceptability. *Language* 87, 249–273. doi: 10.1353/lan.2011.0041

Wilbur, R. (1987). *American Sign Language: Linguistic and Applied Dimensions, 2nd Edn*. Boston, MA: Little Brown.

Wilbur, R. (2002). "Phrase structure in ASL and ÖGS," in *Progress in Sign Language Research in Honor of Sigmund Prillwitz*, eds R. Schulmeister, and H. Reinitzer (Hamburg: Signum), 235–247.

Wilbur, R. (2003). Representations of telicity in ASL. *Chic. Linguist. Soc.* 39, 354–368.

Willis, D. W. E. (1998). *Syntactic Change in Welsh: A Study of the Loss of the Verb-Second*. Oxford: Oxford University Press.

Wojda, P. (2010). "Transmission of Polish sign systems," in *Sign Languages*, ed D. Brentari (Cambridge: Cambridge University Press), 131–148.

Woll, B. (2003). "Modality, universality and the similarities among sign languages: an historical perspective," in *Cross-linguistic Perspectives in Sign Language Research: Selected Papers from TISLR 2000*, eds A. Baker, B. van den Bogaerde, and O. Crasborn (Hamburg: Signum), 17–27.

Wood, S. K. (2007). Degrees of resiliency in acquisition of language. *Nanzan Linguistics; Special Issue* 3, 1, 315–330.

Wood, S. K. (2010). "Acquisition of topicalization in very late learners of Libras," in *Deaf Around the World: The Impact of Language*, eds G. Mathur and D. J. Napoli (Oxford: Oxford University Press), 164–183.

Woodward, J. (1980). "Sociolinguistic research on ASL: an historical perspective," in *Sign Language and the Deaf Community: Essays in Honor of William Stokoe*, eds C. Baker and R. Battison, (Silver Spring, MD: Linstok Press), 117–134.

Woodward, J. (1987). "Sign languages—providence Island," *Gallaudet Encyclopedia of Deaf People and Deafness*, Vol. 3, ed J. V. Van Cleve (New York, NY: McGraw-Hill), 103–104.

Yau, S. C. (2008). "The role of visual space in sign language development," in *Space in Languages of China*, ed D. Xu (Dordrecht: Springer Verlag), 143–174.

Zeshan, U. (2003). Indo-Pakistani Sign Language grammar: a typological outline. *Sign Lang. Stud.* 3, 157–212. doi: 10.1353/sls.2003.0005

Zeshan, U. (2006). "Negative and interrogative structures in Turkish Sign Language (TID)," in *Interrogative and Negative Constructions in Sign Languages*, ed U. Zeshan (Nijmegen: Ishara Press), 128–164.

Zhang, N. N. (2007). Universal 20 and Taiwan Sign Language. *Sign Lang. Linguist.* 10, 55–81. doi: 10.1075/sll.10.1.05zha

Zhang, X., Han, J., and Zhao, G. (2013). "Research on the construction of dynamic chinese natural sign language corpus," in *Mobile Ad-hoc and Sensor Networks (MSN), 2013 IEEE Ninth International Conference on* (Piscataway, NJ; Los Alamitos, CA: The Institute of Electrical and Electronics Engineers, Inc.), 502–505.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

From iconic handshapes to grammatical contrasts: longitudinal evidence from a child homesigner

Marie Coppola[1]* and Diane Brentari[2]

[1] Departments of Psychology and Linguistics, Language Creation Laboratory, University of Connecticut, Storrs, CT, USA
[2] Department of Linguistics, Sign Language Laboratory, University of Chicago, Chicago, IL, USA

Many sign languages display crosslinguistic consistencies in the use of two iconic aspects of handshape, handshape type and finger group complexity. Handshape type is used systematically in form-meaning pairings (morphology): *Handling handshapes* (Handling-HSs), representing how objects are handled, tend to be used to express events with an agent ("hand-as-hand" iconicity), and *Object handshapes* (Object-HSs), representing an object's size/shape, are used more often to express events without an agent ("hand-as-object" iconicity). Second, in the distribution of meaningless properties of form (morphophonology), Object-HSs display higher *finger group complexity* than Handling-HSs. Some adult homesigners, who have not acquired a signed or spoken language and instead use a self-generated gesture system, exhibit these two properties as well. This study illuminates the development over time of both phenomena for one child homesigner, "Julio," age 7;4 (years; months) to 12;8. We elicited descriptions of events with and without agents to determine whether morphophonology and morphosyntax can develop without linguistic input during childhood, and whether these structures develop together or independently. Within the time period studied: (1) Julio used handshape type differently in his responses to vignettes with and without an agent; however, he did not exhibit the same pattern that was found previously in signers, adult homesigners, or gesturers: while he was highly likely to use a Handling-HS for events with an agent (82%), he was less likely to use an Object-HS for non-agentive events (49%); i.e., his productions were heavily biased toward Handling-HSs; (2) Julio exhibited higher finger group complexity in Object- than in Handling-HSs, as in the sign language and adult homesigner groups previously studied; and (3) these two dimensions of language developed independently, with phonological structure showing a sign language-like pattern at an earlier age than morphosyntactic structure. We conclude that iconicity alone is not sufficient to explain the development of linguistic structure in homesign systems. Linguistic input is not required for some aspects of phonological structure to emerge in childhood, and while linguistic input is not required for morphology either, it takes time to emerge in homesign.

Keywords: sign language, homesign, gesture, phonology, morphology, language emergence, iconicity, grammaticalization

INTRODUCTION

Striking cross-linguistic similarities have been described in how sign languages use handshape to mark linguistic distinctions; see, e.g., Brentari et al. (2012) for morphosyntax and Brentari and Eccarius (2010) for phonology[1]. This paper will discuss two aspects of handshapes and explore how these forms are used in a grammatical system longitudinally in a child developing a homesign system. The first is the *handshape type*, which

characterizes the way that a handshape expresses a meaning. Specifically, *Handling* handshapes depict the hand manipulating an object, while *Object* handshapes capture an object's properties by using the handshape to depict the whole item, or size and shape dimensions of the item. Handling handshapes are used to describe events in which an agent manipulates an object, and Object handshapes describe events or arrays of objects that do not involve an agent. Thus, this use of different handshape classes, or types, constitutes a *morphosyntactic* distinction in sign languages. The second dimension of handshape is selected finger group complexity, which involves the selection of phonological groups of fingers. Selected finger groups with higher complexity[2] are associated with

---

[1] For convenience, all of the sign languages mentioned in this paper and their abbreviations are given here: Al-Sayyid Bedouin Sign Language (ABSL), American Sign Language (ASL), Chinese Sign Language (Shanghai dialect; CSL-S), Italian Sign Language (LIS), Israeli Sign Language (ISL), Japanese Sign Language (JSL), Nicaraguan Sign Language (NSL), and Sign Language of the Netherlands (NGT).

[2] We will refer to this as "finger group complexity" henceforth.

Object handshapes, and finger groups with lower complexity are associated with Handling handshapes. This use of a meaningless property of handshape nested within the above-mentioned morphological contrast results in a *morphophonological* distinction in sign languages.

The present study connects to several of the themes of this special issue: specifically, in the way that meaning in natural languages and its phonological vessels (in this case, manual gestures) interact. In particular, iconicity has been proposed as a likely, or even inevitable resource that can be tapped in the processing of sign languages (Vigliocco et al., 2005), in the acquisition of structure in sign languages (Ormel et al., 2009; Thompson et al., 2009), in the emergence of linguistic structure in new languages (Meir et al., 2007), in the very organization of sign language grammars (Cuxac, 1999; Demey and van der Kooij, 2008; Meir, 2010), and as a general property of both signed and spoken languages (Perniss et al., 2010). In the realm of experimental semiotics, Fay et al. (2013, 2014) argue that gesture is likely to bootstrap human communication systems in the absence of linguistic input precisely because it affords greater iconicity than the auditory modality. We will argue that iconicity is a multilayered, complex notion that must be treated with care, especially when evaluating its influence on the distribution of linguistic components in grammatical systems.

Accessibility to iconicity in development does not happen all at once in sign language or gesture. For example, Brentari et al. (2012) have described *hand-as-hand* iconicity in Handling handshapes as distinct from *hand-as-object* iconicity in Object handshapes when gesturers and users of conventional sign languages describe events; Padden et al. (2013) apply a similar distinction in their work as well. The level of accessibility to different kinds of iconicity depends on the ambient language, the age and life experience of the participant, as well as the nature of the task. The handling of objects is a human action, argued to be easier to produce in gesture than a static object or action by an object (Piaget, 1952; Werner and Kaplan, 1963). On the other hand, studies of child gesture show that Object handshapes are used before Handling handshapes (Kaplan, 1968; Overton and Jackson, 1973; Boyatzis and Watson, 1993; O'Reilly, 1995; Tomasello et al., 1999). Namy et al. (2004) argue that iconicity as a factor in concept acquisition is not immediately available to infants and toddlers, and takes time to learn. Moreover, in all of these studies one must also ask whether the experimental task itself was designed to address specific minimal differences in meaning that can be expressed in componential form; for example, the two properties of handshape relevant here are elicited using a vignette description task which targets minimal differences in handshape meaning, while carefully controlling for location and movement.

Both morphosyntax and morphophonology use iconicity in different and independent ways, which will be described in the next two sections. Moreover, while the importance of iconicity as a source of raw material for new forms in sign languages cannot be ignored, the *distribution* of these elements is abstract and can also be arbitrary: *handshape type* pertains to morphosyntactic representation while *selected finger group complexity* is at the level of morphophonological representation (Brentari, 2007, 2011). We begin with an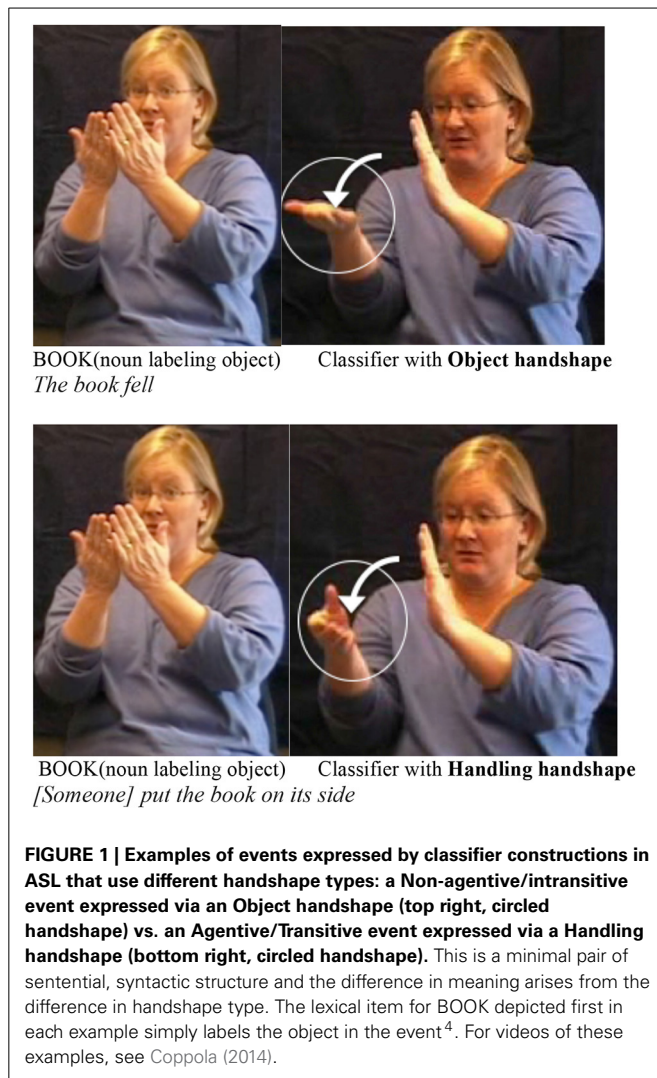 overview of how handshape type is organized in the morphosyntax of sign languages, and then describe how finger group complexity varies systematically within those morphological handshape classes, and gives rise to morphophonological structure. In two studies, we examine the development of the uses of these aspects of handshape in one homesigning child, and use comparative analyses with other participant groups to uncover the sources of convergence between his patterns and the crosslinguistic patterns we have observed in sign languages.

MORPHOSYNTAX IN CLASSIFIER CONSTRUCTIONS

Despite the apparent iconicity of Object handshapes and Handling handshapes, both types of handshapes can contribute to morphological structure. Events involving the motion and location of people and objects are preferentially described in sign languages using handshapes in particular configurations and orientations, combined with movements and locations[3]. These morphologically complex predicates are known as *classifier constructions* in the sign language literature (see Emmorey, 2003 and Zwitserlood, 2012 for summaries and examples from a variety of sign languages), and these components have been analyzed as discrete, meaningful, productive forms that are stable across related contexts (Supalla, 1982; Emmorey and Herzig, 2003; Eccarius, 2008). Our work pertains to "whole entity" and "SASS" (descriptor) classifiers (considered Object classifiers), and to "Handling" classifiers, exemplified in **Figure 1**. The Object handshape circled in **Figure 1** (top) uses hand-as-object iconicity to represent a flat object, the book itself, involved in a "falling" event. The Handling handshape circled in in **Figure 1** (bottom) uses hand-as-hand iconicity to represent someone holding/moving a book.

Benedicto and Brentari (2004) found that handshape types (Handling vs. Object) are not simply morphological forms with discrete meanings, but rather check features of argument structure within the syntactic tree. The "willing" test and the "imperative" test are widely known to detect agents crosslinguistically (Van Valin, 1993); by adding the sign WILLING or FINISH (with the negative imperative meaning "stop!") to ASL sentences containing the classifier predicates in **Figure 1**, it is possible to detect the presence of an agent. Both WILLING and FINISH can be added to the sentence in **Figure 1** (bottom) to obtain well-formed, grammatical sentences (English translation from ASL: "*[Someone] willingly put the book on its side*" and "*Stop putting the book on its side!*"). In contrast, adding WILLING or FINISH to the sentence in **Figure 1** (top) obtains ungrammatical sentences (English translation from ASL: *"The book fell willingly"* and *"Book, stop falling!"*). Since the only part of the structures in **Figure 1** that varies is the handshape, the differences obtained using these diagnostic syntactic tests is attributed to handshape, indicating that the sentence with the Handling handshape (**Figure 1**, bottom) has an agent, while the one with the

---

[3]As one reviewer suggested, it is an open question whether a system will express one meaning with both handshape and movement fused together or whether handshape and movement will have componential meanings. Evaluating this hypothesis will require a parallel and systematic analysis of movement, similar to the one for handshape presented here. This is beyond the scope of this paper, but is an active area of our current work.

BOOK(noun labeling object)   Classifier with **Object handshape**
The book fell

BOOK(noun labeling object)   Classifier with **Handling handshape**
[Someone] put the book on its side

FIGURE 1 | Examples of events expressed by classifier constructions in ASL that use different handshape types: a Non-agentive/intransitive event expressed via an Object handshape (top right, circled handshape) vs. an Agentive/Transitive event expressed via a Handling handshape (bottom right, circled handshape). This is a minimal pair of sentential, syntactic structure and the difference in meaning arises from the difference in handshape type. The lexical item for BOOK depicted first in each example simply labels the object in the event[4]. For videos of these examples, see Coppola (2014).

FIGURE 2 | Position of finger selection (circled) within sign language phonology feature tree for handshape. Reprinted with kind permission from Springer Science+Business Media B.V.

MORPHOPHONOLOGY IN CLASSIFIER CONSTRUCTIONS

The hand is not treated as an undifferentiated whole in sign language phonology; handshape has several sub-components. The representation of handshape includes a branch in the feature tree representing the "active," or *selected fingers* in a given handshape (see **Figure 2**). Selected fingers are those that move or contact the body during the articulation of a sign. Contrasts in selected fingers constitute minimal pairs and are important for the application of phonological rules in several sign languages, including ASL, ISL, and NGT (Klima and Bellugi, 1979; van der Hulst, 1995; Meir and Sandler, 2007). The features of *selected fingers* form natural classes of handshapes, such as the *index finger group* ⟨⟩, which contains ⟨⟩, ⟨⟩, ⟨⟩, ⟨⟩, and ⟨⟩; that is, all handshapes with only the index finger selected are a natural class of handshapes, similar to the grouping of obstruents in English[7].

The morphological categories for Object and Handling handshapes in classifier constructions might or might not be paralleled by a corresponding phonological pattern. Using joint configuration as an example, let us consider a hypothetical situation in which a given sign language were to use the following set of handshapes as whole entity classifiers—⟨⟩ ⟨⟩ ⟨⟩ ⟨⟩ ⟨⟩ ⟨⟩. This set would not only be a morphological class, but would also form a phonological class, because the selected fingers in each handshape share a phonological property; namely, they are all *fully open* ("extended"). If a signer of this sign language were to encounter new handshapes, such as ⟨⟩ ⟨⟩ or ⟨⟩, these would be predicted to belong to the whole entity morphological class because of this phonological generalization. In contrast, if a second hypothetical sign language were to use this set of handshapes for whole entity classifiers—⟨⟩ ⟨⟩ ⟨⟩ ⟨⟩ ⟨⟩ ⟨⟩ ⟨⟩— the set could still be a morphological class, but it would not form a phonological class, as there is no common joint property that the handshapes share. The handshapes would constitute a morphological, but not a

Object handshape (**Figure 1**, top) does not. The sensitivity of classifier handshape types to such tests is evidence that they are part of the morphosyntax[5]. Besides ASL, adult users of Italian Sign Language (LIS) (Mazzoni, 2009) and NSL also employ this pattern (Goldin-Meadow et al., under review)[6], as do deaf children acquiring ASL and LIS, but it takes time to develop in children (Brentari et al., 2013, in press).

---

[4]To be as neutral as possible about the linguistic status of such gestures in the different groups we will be comparing, we refer to these as "labels." Iconic handshapes have a different distribution in labels/nouns, and because we are comparing the portions of participants' responses that are most comparable to sign language classifier constructions, we exclude the labels from the analyses in the current study.

[5]This phenomenon is phrasal because it is the combination of lexical semantic features that causes the ungrammaticality. So it can be analyzed as either phrasal semantics, or phrasal syntax. Under a syntactic model that allows the features of the lexical semantics to be fed into the syntactic tree for feature checking (e.g., Borer, 2005) it is considered part of syntax.

[6]While many studies have investigated classifier constructions cross-linguistically (e.g., Aronoff et al., 2003) and have compared signers with non-signing gesturers (e.g., Schembri et al., 2005), a relatively small number
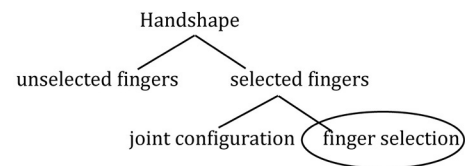

(in addition to those already cited) have examined this morphosyntactic distinction in ASL and other sign languages (Schick, 1987; Kegl, 1990; Janis, 1992; Brentari et al., 2001; Zwitserlood, 2003; and Pfau and Steinbach, 2006).

[7]We use the handshape font (e.g., ⟨⟩) to indicate handshapes. When used for individual handshapes, the image is a picture of a particular hand. When this font is used to represent finger groups, the image stands for a category of handshapes. In these cases, the term "finger group" will precede the image (i.e., "finger group ⟨⟩"), and the image will picture a handshape with extended fingers, without the thumb. For example, the finger group ⟨⟩ represents the set of handshapes that includes the range of configurations with the index and middle finger selected: ⟨⟩, ⟨⟩, ⟨⟩, ⟨⟩, ⟨⟩, ⟨⟩, and ⟨⟩.

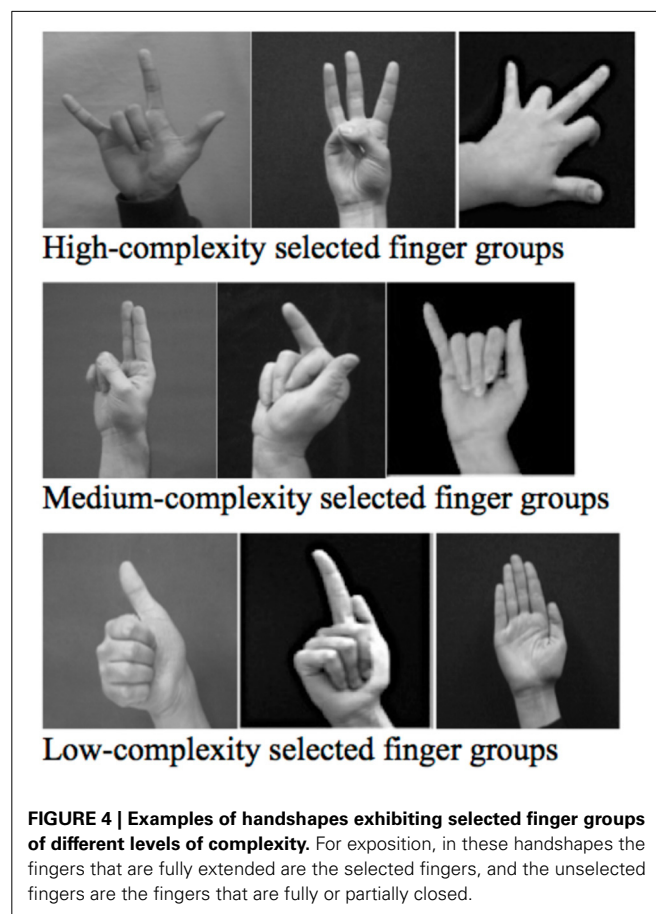phonological, class (see **Figure 3**); that is, the handshapes in this class would mark a specific class of meanings (whole entities), but without a corresponding phonological property that unifies them. In this second case, a signer could not predict from the handshape structure the morphological class to which it belongs.

Returning to the current study, across a number of sign languages, Object handshapes used in classifier constructions display higher finger group complexity on average than Handling handshapes: ASL and LIS (Brentari et al., 2012), as well as in CSL-S and NSL, two sign languages unrelated to ASL or LIS, and in children acquiring ASL, LIS and NSL after 4–6 years of exposure (Brentari et al., in preparation)[8]. The Object and Handling classifier handshapes in these sign languages therefore exhibit not only morphosyntactic structure but also morphophonological structure of the sort described above—relatively high average finger group complexity associated with Object handshapes and relatively low average finger group complexity associated with Handling handshapes. This distinction in complexity has been described as indirectly iconic: finger group selection is associated with representing the physical properties of objects, and joint configuration associated with manipulating objects, because each aspect is iconically adapted for that phonological task (Brentari et al., in preparation).

RANKING THE FINGER GROUP COMPLEXITY OF HANDSHAPES IN SIGN LANGUAGES

Finger group complexity is based on a number of factors, including frequency (Hara, 2003; Eccarius and Brentari, 2008), age acquired (Boyes Braem, 1990), and the number of branches in the phonological structure (Brentari, 1998). Higher finger group complexity also indicates a larger and more mature inventory of handshapes (Marentette and Mayberry, 2000). Handshapes can be divided into three levels of *finger group complexity* based on these criteria. Handshapes with *Low finger group complexity* (**Figure 4**, bottom) have the simplest phonological representation (Brentari, 1998), are the most frequent cross-linguistically (Hara,

---

[8]Handshape complexity can of course be measured along a variety of dimensions, including joint complexity. We focus our attention here on selected finger complexity because the pattern of joint complexity (higher for Handling-HSs than for Object-HSs) does not differ between signers and silent gesturers (Brentari et al., 2012, in preparation), and is therefore not a likely place to observe the development of linguistic structure in emerging languages.

2003; Eccarius and Brentari, 2008), and are the earliest acquired by native signers (Boyes Braem, 1990). They include the groups with all fingers (finger group), the index finger (finger group), and the thumb (finger group). Low-complexity handshapes account for an overwhelming proportion of handshapes in NGT, JSL, and ASL (81% in ASL from Hara, 2003). *Medium finger group complexity* handshapes (**Figure 4**, middle) include one additional structural elaboration: either a second selected finger on the radial (thumb) side of the hand (the default side), as in finger group , or a single selected digit that is not on the radial side, as in the

FIGURE 4 | Examples of handshapes exhibiting selected finger groups of different levels of complexity. For exposition, in these handshapes the fingers that are fully extended are the selected fingers, and the unselected fingers are the fingers that are fully or partially closed.

whole entity classifiers—language 'x'
✔ **Morphological class**
✔ **Phonological class**
  All handshapes have a fully
  "open" joint specification
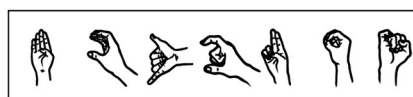
whole entity classifiers—language 'y
✔ **Morphological class**
∗ **not a Phonological class**
  joint specifications
  differ across the handshapes

FIGURE 3 | A group of hypothetical handshapes that would constitute a morphological and phonological class (left) and another group that would constitute only a morphological class (right).

pinky or middle finger groups 🖐 and 🖐. *High finger group complexity* handshapes (**Figure 4**, top) include all remaining finger groups, which are less frequent and have more complex phonological structures. The correspondence of different levels of finger group complexity to different handshape types is evidence for the two levels of grammar; this organization is not evident in the silent gestures of hearing people (Brentari et al., 2012, in preparation). It is not duality of patterning[9] because the two levels of structure are not independent, and it is not absolute[10], but importantly, it reflects organization of the grammatical system at two levels: morphological and phonological.

In summary, iconicity, syntax, morphology, and phonology are complex components of a sign language, and each is acquired along a unique time course. Moreover, the principles of each component interact with one another via interface principles and constraints, as we have seen above: the agent/non-agent distinction (syntax) and the contrast in finger group complexity (phonology) can become manifest only after a distinction between Object and Handling handshapes exists (morphology). Here we examine for the first time the use of Handling and Object handshapes longitudinally in one child homesigner in order to address how such morphosyntactic and morphophonological patterns might emerge.

HOW MIGHT THESE PATTERNS ARISE IN HOMESIGN?

How might these systematic uses of handshape type and finger group complexity have arisen independently in these unrelated sign languages? Sign languages have their roots in homesign systems, which are gesture systems created by individuals in the absence of a conventional language model (Coppola and Senghas, 2010; Brentari and Coppola, 2012); homesign systems, in turn, use as their raw materials the gestures produced by hearing people in the surrounding culture (Fusellier-Souza, 2006).

A homesigner is a deaf individual whose degree of deafness prevents sufficient access to spoken language to permit acquisition (Goldin-Meadow, 2003). This lack of access to spoken language structure, or to formal instruction, precludes homesigners' learning to read and write. Homesigners also have no or extremely limited access to and interactions with other deaf people, especially to signers of a sign language (either an established sign language, such as ASL for homesigners in the United States, or an emerging sign language, as is used by members of the Deaf community in Nicaragua). Homesigners do not interact regularly with other deaf people, and are not members of a Deaf community[11].

With regard to the morphosyntactic distinction, adult homesigners as a group behave similarly to users of sign languages and exhibit the agentive/non-agentive distinction (Goldin-Meadow et al., under review). Regarding silent gesture, (Brentari et al., in press) found that hearing gesturers as a group do not use Object and Handling handshape types systematically to express agentivity, and there is considerable between-subject variation; some individual gesturers can produce this pattern (notably adult, Italian gesturers). This contrastive use of handshape is unlikely to be due to the presence of a grammar, as in sign languages. Rather they argue that language, culture, and cognition, as well as the task, contribute to the gesturers' performance. Gesturers are asked to describe minimally contrastive vignettes—they see exactly the same object in the same situation with the minimal difference being the presence of an agent. They are using silent gesture, and therefore channeling all communicative energy into the manual modality. They have a spoken language, and therefore have had a model for the type of meaningful contrast being elicited. Italians also live in a culture that uses a large number of emblematic gestures, which may also provide an additional advantage.

With regard to the morphophonology—higher average finger group complexity in Object-HSs than in Handling-HSs—Brentari et al. (2012) found that in adult homesign systems Object handshape finger group complexity was as high as that of ASL and LIS signers, while homesigners' Handling handshape complexity tended to be higher than signers'. Gesturers do not produce the morphophonological pattern observed in adult and child signers; indeed, as described above, they exhibit large individual differences[12].

Little is known about the development of these aspects of linguistic structure in children who do not receive conventional linguistic input. In the current study, we will directly compare four adult homesigners in Nicaragua with a child homesigner in Nicaragua, called Julio, performing the same task over time, coded and analyzed in the same way as the adults. Julio stands at the intersection of three different types of populations/participants: (1) as a homesigner, we can compare him to the previously studied adult homesigners; (2) as an individual whose resources for expression in the manual modality are limited to the visual aspects of the language and communication in his environment, we can compare him to hearing, non-signing individuals using silent gesture[13]; and (3) as a child, we can compare him to other children, who are in similar developmental stages, but have different linguistic backgrounds—Deaf children acquiring a sign language from signing parents, and hearing children acquiring spoken language.

---

[9]Duality of patterning is not synonymous with phonological organization. Here, the organization of finger group complexity (a type of phonological organization) is isomorphic with the morphological organization, and thus does not show an independent level of organization (i.e., duality of patterning).

[10]This is based on average language-specific, finger group complexity for handshape type; every Object-HS is not more complex than every Handling-HS.

[11]More details about how and to what extent homesigners communicate with the hearing people around them are provided in the descriptions of the child and adult homesigners in this study in the Participants Section.

---

[12]In two previous studies, hearing gesturers as a group either showed the opposite pattern—higher finger group complexity in Handling handshapes (Brentari et al., 2012)—or very little finger group complexity overall, with no difference between Handling and Object handshapes (Brentari et al., in preparation).

[13]Although unlike Julio, the hearing children have linguistic input and an existing grammatical system, just one that does not use handshape for grammatical functions.

FROM GESTURE TO GRAMMAR: A DISTRIBUTIONAL MODEL

One of the major advantages of our comparative approach is that it allows us to disentangle the contributions of several factors to the emergence and development of linguistic structure: the presence and quality of linguistic input; developmental stage; the function of the gestures/signs in an individual's life (as a primary language for homesigners and signers vs. a one-time occurrence for hearing gesturers); and culture (though this is not directly addressed by the present studies). Because Julio is a child homesigner, we have the unique opportunity to track the changes in his homesign system as it acquires more linguistic structure, from its roots in the non-linguistic gestures produced by the hearing individuals around him. Specifically, we use the distribution of Object and Handling handshapes as a metric of the linguistification of his homesign system[14]. As described previously, this structure is manifested in the distribution of handshapes by exploiting *hand-as-hand* iconicity to use Handling-HSs to express events in which an agent manipulates an object, and exploiting *hand-as-object* iconicity to use Object-HSs for events without an agent. We propose the following model of the emergence of systematic distributions of these two aspects of handshape in the absence of a linguistic model. While these stages are stated in terms of how these aspects of handshape might be selected from the raw materials available in gesture and shaped into linguistic elements, in principle these stages apply to any aspect of form that undergoes the process of transformation from a gesture into a linguistic element *via systematic re-organization and distribution of iconic properties*.

Stage 1: Recognizing and using Handshape Type (i.e., Object- and Handling-HSs) as an aspect of form that can be utilized for a meaningful contrast to describe events involving objects and their manipulation. Using the hand in other ways—for example, to trace the path of an object as it moves through space, or as a mere extension of the arm—exemplified by a child who flaps his or her arms to represent an airplane flying—do not reflect this recognition of the affordances of handshape type, and are therefore unlikely to show further development of handshape as a marker of linguistic contrasts. This stage thus represents a potential stage that is not observed when children acquire these grammatical subsystems from linguistic input.

Stage 2: Distinguishing between classes of Object and Handling handshapes in one's system; one manifestation of this would be to associate one handshape type (e.g., Object-HSs) with one event type (Non-agentive events), and the other handshape type (e.g., Handling-HSs) with Agentive events. However, this association does not have to be complete in order for these two handshape classes to emerge.

Stage 3: Phonological organization that mirrors the morphological organization, but is not necessarily independent from it, i.e., higher finger group complexity in Object-HSs than

in Handling-HSs, as shown in **Figure 3**. Note that difference in the distribution of Object and Handling handshapes is all that is needed before this **morphophonological** pattern can develop.

Stage 4: Using Handshape Type to mark a linguistic contrast— i.e., using one handshape type for one purpose and the other handshape type for the other. Specifically, we see the complete association of non-agentive events with Object-HSs (and their on average higher finger group complexity) and agentive events with Handling-HSs, as is observed in the **morphosyntax** of classifier constructions in established sign languages.

Our broad research questions, then, center on how these linguistic uses of handshape develop over time in an individual, in the absence of conventional linguistic input, and their relative timing of emergence. Specifically, in Study 1 we follow the trajectory of the morphosyntactic distinction of handshape type during the 5-year period ending when Julio was about 12½ years old. In Study 2 we follow the trajectory of the morphophonological distinction involving finger group complexity over the same period.

Prior work with child signers suggests that they follow stages 2–4 and may be able to breeze through Stage 1 because they see people signing around them. If Julio behaves like them, his systematic distribution of finger group complexity will precede the sign-like agentive/non-agentive contrast of handshape type (i.e., phonology before morphology). However, if he behaves more like an adult gesturer who is using the resources of iconicity and world experience and who retains as much iconicity as possible in order to better communicate with communication partners who are not skilled users, we would see evidence of the agentive/non-agentive pattern of handshape type earlier than a morphophonological pattern, where the iconicity is less available (i.e., morphology before phonology).

STUDY 1: MORPHOSYNTAX

Supalla (1982) first examined handshape in the acquisition of sign language event descriptions, but did not focus on the agentive/non-agentive opposition. Schick (1987) elicited handling and object classifiers (handshapes) in event descriptions from 24 ASL-learning children (ages 4;5–9;0). Although the children generally used Handling-HSs and Object-HSs correctly, at every age they were more likely to produce correct Object-HSs than Handling-HSs, demonstrating an Object-HS bias (Table 5.3, p. 78). Slobin et al. (2003) found that children learning ASL and NGT spontaneously produced Handling-HSs and Object-HSs as early as age 2;5, but their analysis did not differentiate handshapes used to label objects and actions (e.g., nouns and verbs) from productive classifier predicates (event descriptions, in the terms used here), so it was not possible to determine whether the handshape appeared in a classifier predicate or in a lexical item produced (or created) by the child. Brentari et al. (2013) studied noun/label and classifier/event description Handling-HSs and Object-HSs separately, and found, like Schick (1987), that children produced Handling-HSs for events with an agent less consistently than they produced Object-HSs for events without an agent, and that the agentive/non-agentive opposition is mastered in Deaf

---

[14]Because we are studying gesture systems in which we cannot assume the existence of a classifier system a priori, henceforth we use the more neutral term "handshape" and refer specifically to Object handshapes (Object-HS) and Handling handshapes (Handling-HS). Likewise, we use the neutral term "event description" for constructions that parallel, in form and function, the classifier predicates of sign languages.

children acquiring ASL from native input between 7 and 10 years of age[15].

Importantly, before handshape can be successfully used to encode a morphosyntactic opposition as described in Stage 1, the child must be producing meaningful (iconic) handshapes. The hands can express events in many other ways, for example, by using an index finger or the whole hand to simply trace the outline or the path of an object that is placed on, or picked up from, the table. In such cases, the movement is iconic but the handshape is not. Other attested examples include using the whole body to represent the movement of an object, such as a child extending her arms to depict an airplane and then leaning over to indicate that the airplane is falling. Both Deaf signing children and child and adult hearing gesturers (Brentari et al., in press), produce such forms, though they are produced much more frequently by non-signers (gesturers asked to respond using only their hands).

We can neither perform traditional syntactic tests that rely on grammaticality judgments (as described in the introduction), nor do we expect minimal pairs or phonological assimilation rules among homesigners and gesturers. However, we have developed methods for identifying such patterns within such systems, if they exist, by comparing the results from such diagnostics with the distribution of handshapes in elicited productions (Brentari et al., 2012). We describe these methods in more detail below.

We focus here on the forms that are comparable to the classifier constructions used by signers; namely, those used in the Event descriptions (i.e., label responses that identified the object were not included)[16]. Here we will ask whether, by the age of 12;8, the child homesigner uses Handling handshapes to express events with agents, and Object handshapes to express events without agents, as in the sign language pattern previously identified. If so, when does this pattern emerge? How does it compare to the patterns produced by adult homesigners, child signers, and child gesturers? To address these questions, we will compare the handshape patterns produced by Julio across the sessions, and then compare his performance with the patterns exhibited by each adult homesigner previously studied in Nicaragua, using the same stimuli, procedure, coding, and analyses. We then ask whether Julio's rate of iconic handshape production (Handling-HSs and Object-HSs together) more closely resembles that of adult homesigners, or that of signing or gesturing children from previously published work (Brentari et al., 2013, in press).

METHODS
Participants
Child homesigner. The new data reported here were collected from one deaf homesigning child in Nicaragua called "Julio," who was tested at five time points between the ages of 7;4 (years; months) and 12;8. Julio's family reports that Julio has been deaf since birth. Though audiometry results were not available, his degree of deafness has prevented him from acquiring spoken Spanish. In addition, during the period of the study, Julio did not have sufficient exposure to Nicaraguan Sign Language (NSL) to acquire it. At the final testing session, he began to demonstrate very limited use of NSL signs. The first author became acquainted with Julio through outreach procedures conducted by the Center for Special Education in Estelí, Nicaragua, a medium-sized city about an hour and a half north of the capital, Managua. Julio and his family live in a relatively poor area; his family (all hearing and non-signing) is not very engaged with him, with the exception of his grandmother, who does not gesture extensively with him. Despite repeated visits by first author and the outreach coordinator over the 5-year period of the study to emphasize its importance, Julio's school attendance was sporadic at best. According to his teacher, he did not attend school at all between the ages of 9 and 11[17].

Because Julio was attending school so sporadically, we observed very little influence of NSL on his homesigns. He acquired very few lexical items, even highly frequent ones: for example, he did not even acquire the NSL count list, a set of signs used often in the classroom, during the period under study. The lack of use of highly frequent lexical items (e.g., *man, woman*) and routine phrases (e.g., *good morning*), combined with our observations of his gesturing with the other deaf children in the classroom who were acquiring NSL, led us to conclude that he was not receiving sufficient exposure to NSL to acquire it during the period of our study. He had no communication partners who used NSL with him outside of the school setting. His brother, who is 2 years older, is his main homesign communication partner; however, recent research examining lexical conventionalization (Richie et al., 2014) and grammatical structure (Carrigan and Coppola, 2012, in preparation) suggests that regular interactions using homesign do not guarantee shared structure between the homesigner and his or her hearing family members.

Adult homesigners[18]. Four adult deaf homesigners (1 female) living in Nicaragua also participated in the study (mean age 24, range 20–29 years). The adult homesigners had no congenital cognitive deficits, had not learned spoken or written Spanish, and had not acquired NSL. None had attended school regularly. The adult homesigners did not interact with one another and each had developed a homesign system of his or her own that, unlike Julio, they continued to use as their primary language into adulthood (Coppola and Newport, 2005) and which exhibit a range of linguistic properties, such as pronouns (Coppola and Senghas, 2010) and devices expressing quantity akin to plurals (Coppola et al., 2013). As was the case for the child homesigner for the majority of our study period, the adult homesigners use

---

[15]Other studies of the acquisition of classifiers in sign languages include Bernardino, 2006; Tang et al., 2007; however, these studies also do not take up the question of how children begin to use handshapes differently in agentive vs. non-agentive events.

[16]A discussion of the time course of the development of labels is beyond the scope of this paper, but is an active topic of our current work; see Goldin-Meadow et al. (under review) for such a discussion in Nicaraguan adult signers and homesigners.

[17]For two years, there were two young deaf sisters who used Nicaraguan Sign Language at home with their older brother and sister who are Deaf; however, Julio did not attend school regularly during this time.

[18]A subset of these data have been previously published in (Brentari et al., 2012). The analyses reported here replicate and extend those findings by expanding the dataset.

homesign exclusively to communicate with the hearing people around them, and these hearing individuals often communicate with the homesigner using gestures, though communicative success varies greatly across family groups (Carrigan and Coppola, 2012, in preparation). Each homesigner works, makes money, and interacts socially with hearing friends and family, but is not a member of a Deaf community, and does not have regular NSL communication partners. The first author has worked with three of the adult homesigners since 1996, and the fourth since 2004.

Child groups. The analyses in Use of iconic and non-iconic handshape types across groups and Analysis of Specific Handshapes situate the distribution of Julio's handshape types with those produced by child signers and child gesturers (data from Brentari et al., in press). The children in these studies were Deaf native signing children and hearing, gesturing children with no exposure to a sign language responding with silent gestures (3 ASL, 4 LIS, age 3;10–6;4, mean = 5;2); 3 American, and 4 Italian child gesturers (ages 4;3–5;3, mean = 4;8)[19].

Stimuli

The stimuli consisted of 118 photographs and short videos (henceforth vignettes)[20]. Eleven object types were used in the vignettes: airplane, book, cigar, lollipop, marble, pen, string, tape, television set, and tweezers. The actual objects depicted in the stimulus clips exhibited a range of colors, shapes, and sizes. Each object type was portrayed in 10 variations that fell into two types of events: (1) Five Non-Agentive events, which depicted a stationary object or an object moving on its own without an agent, and (2) Five Agentive events, which depicted an object being moved by the hand of a human agent[21] (**Figure 5**). Supplementary Material displays the items presented at each testing session.

Procedure

The first author showed each stimulus event to the participant on a laptop computer and elicited a description using minimally verbal instructions: by producing a quizzical facial expression, shrug, and manual flip gesture, often combined with a point. For all sessions, the child homesigner responded to the experimenter (the first author), who has worked with him since he was

Scenes without an agent (Non-agentive events)	Scenes with an agent (Agentive events)
1. [object] on table	6. Put [object] on table
2. [object] on table upside down	7. Put [object] on table upside down
3. *Multiple [objects] on table (regular arrangement in row/s)*	8. *Put multiple [objects] on table (regular arrangement in row/s)*
4. Multiple [objects] on table (random arrangement)	9. Put multiple [objects] on table (random arrangement)
5. [object] moving without an agent (typically falling)	10. Demonstrate function of [object]

Variation 3: Airplanes in a row (static image)   *Variation 8: Put airplanes in a row (video)*

FIGURE 5 | Descriptions of Non-Agentive and Agentive events, variations in number and arrangement of objects, and specific examples of two stimulus events that contrast only in the presence of an agent. Reprinted with kind permission from Springer Science+Business Media B.V.

6;4, and is very familiar with his gesture system[22]. This procedure successfully elicited gestured descriptions from Julio. The adult participants produced their responses to a family member or friend who was familiar with their homesign: Adult 1 (friend), Adults 2 and 4 (siblings), and Adult 3 (mother).

These descriptions were video recorded, transcoded, and clipped into individual files, one file for each vignette description. The responses were transcribed using ELAN (Crasborn and Sloetjes, 2008; ELAN), a tool developed for multimodal language analysis at the Max Planck Institute for Psycholinguistics in Nijmegen, The Netherlands.

Coding

Coding different components of the response. We divided the descriptions that Julio produced for each vignette into two portions: *labels* referring to the objects and *descriptions of the event* depicted in each vignette[23] (see **Figure 6**). Because all of the vignettes in our study show items on a table or being put on a table, we were able to use a sign's location and orientation to categorize it as an *object label* or *event description*. If the participant produced a form that depicted the movement or arrangement in the vignette, the sign was considered an event description; these were typically produced in a specific location within a single plane, or in relation to a secondary object, most often in the horizontal plane of the signing space (reflecting the fact that the objects in our stimuli

---

[19]The interlocutor was a native speaker or signer in all but 3 of 14 sessions, and data were collected, according to the parents' preference, at the child's school, in the child's home, or at the regional headquarters of the Deaf Association in Milan (ENS).

[20]Some of the 10 variations of number and orientation described in **Figure 4** were represented by multiple trials, thus the overall total number of vignettes was higher than 110 [11 (objects) × 10 (variations)]. For example, variation #5 for the object airplane had two versions of movement without an Agent: one in which the airplane fell off the edge of a table, and one in which the airplane (a wind-up mechanical toy) tumbled over itself. For some objects (e.g., pen), variations #8 and 9 showed the agent picking up the objects (vs. putting them down).

[21]The Agentive events in our stimulus set all involve manipulation of an object by a human agent. Objects can of course be manipulated by other objects, for example, tools. In describing such an event, in which a tool is manipulated by a human agent, it would most likely be represented by signers using a Handling handshape in our classification system.

[22]Julio's brother, who is his primary communication partner using the homesign, was usually in school during our testing sessions with Julio and was therefore unavailable to serve as his interlocutor.

[23]Gestures that were not labels or event descriptions, such as gestures indicating the number of objects, were classified as "Extra Information" and were not coded further.

FIGURE 6 | Examples of two complete responses produced by the child homesigner (for videos, see Coppola, 2014**).** The coding and analysis categories, and the values that were assigned to each form, appear below each photo. All analyses reported in Studies 1 and 2 were conducted only on the signs produced in the Event Description portion of each response. Object handshape (top): Stills taken from a response that was produced to describe the stimulus event *"A mechanical toy airplane tumbles over itself repeatedly."*; Handling handshape (bottom): Stills taken from a response to the stimulus event *"Someone puts a toy airplane on a table."* Responses from the adult homesigners were treated identically.

were placed on a table). If the participant's gesture was produced on the body or at a nonspecific location in one of the three planes of neutral space[24], it was considered a label for an object[25].

Coding handshape type: Object vs. Handling handshapes. We categorized each handshape according to type: (1) *Object* handshapes captured properties of the object they represented, either the whole item or size and shape dimensions of the item, and (2) *Handling* handshapes captured properties of the hand manipulating the object. A response was coded as *Both* if: a Handling- and Object-HS simultaneously represented the event on each hand (e.g., a "C" Handling-HS on one hand holding a "B" Object-HS on the other); or the handshape started as a Handling-HS and ended as an Object-HS (e.g., "C" changed to "B") or vice-versa.

  Both responses accounted for 6% of the data for events with an Agent and 5% for No-Agent events. In addition, the following handshapes were classified as *Other*: handshapes that "traced" the outline of the object or the path that it took in the vignette (e.g.,

an index finger or neutral handshape). Handshapes that were neither Handling nor Object handshapes comprised 8% of all productions; 6% for Agent events, and 14% for No-Agent events.

Reliability. Two coders transcribed and coded the child homesign productions. They both coded the same subset of 22 items (54 gestures produced overall) to establish reliability of the coding categories. Inter-rater agreement for classifying a given gesture as part of the object label, event description, or other information was 93%; for classification of handshapes according to configuration, 91%, and for classifying handshapes according to type, as *Object, Handling,* or *Other,* 88%[26]. Discrepancies were resolved through discussion.

RESULTS: AGENTIVE/NON-AGENTIVE DISTINCTION USING HANDSHAPE

All of the homesigners produced at least one response per vignette[27]. We first present the results of the longitudinal analysis

---

[24] There are three planes in the signing space: the horizontal plane, the vertical plane, and the mid-sagittal plane (Brentari, 1998).

[25] These criteria for categorizing a particular gesture or sign according to its function as a label or as an event description are reliable in the context of this task, and we make no claim that these criteria would be appropriate to distinguish nouns and verbs across an entire sign language.

[26] For the adult homesign data, interrater reliability for classification of handshapes according to configuration and for classifying handshapes according to type was 90% (Brentari et al., 2012). (Labels were not coded for those analyses).

[27] Three of Julio's original, complete videotaped responses have been annotated and may be viewed online (see Coppola, 2014).

examining Julio's responses over time (all 11 objects, all conditions, all handshape types), followed by a comparison with the four previously studied adult homesigners in Nicaragua for 4 objects (airplane, book, lollipop, and pen)[28], followed by a comparison with child signers and gestures on two objects (airplane and lollipop), plus the variation in which an object moved without an agent (e.g., it fell) (all 11 objects).

Longitudinal analysis

Julio used handshape contrastively with respect to the presence of an agent from the earliest age studied, although not in the pattern seen in previous studies with adult homesigners or signers of established sign languages. Chi-square tests revealed a significant association between Handshape Type and the presence of an Agent for four of the five testing sessions (**Table 1**). Across all sessions, Julio, like signers of established sign languages, produced Handling-HSs for events with an Agent ($n = 258$; mean = 82%, black bars, **Figure 7**, right chart); however, for events with No Agent ($n = 179$), he produced the expected Object-HS on average only 49% of the time (gray bars, **Figure 7**, left chart). We

[28] These four objects represent a range of object properties with respect to shape, size and complexity of features, which we expected would affect the handshape configurations that would be used to describe them. For this reason, they were also the objects that were already analyzed for the adult homesigners and thus available for comparison.

chose this analytical approach because our primary interest lies in the association between Handshape Type and the presence of an agent (assessed by the Chi-square test), rather than in the relationship between the proportions of Object- and Handling-HSs produced within a particular session. Accordingly, the results in **Figure 7** are organized to highlight the different distributions of Handshape Type in No-Agent vs. Agent events[29].

Julio compared with adult homesigners

Like Julio, the responses of all four adult homesigners in Nicaragua demonstrated a significant association between handshape type (Object vs. Handling) and the presence/absence of an Agent (**Table 2** summarizes the Chi-square analyses). However, Julio's responses to events with an Agent differed from those of the adult homesigners in terms of both pattern type and consistency. **Figure 8** shows the proportions of Object-HSs and Handling-HSs produced in each context by Julio and the four adult homesigners. Julio and Adults 1 and 4 preferred Handling-HSs for Agentive events; However, Adult 2 was more likely to produce Object-HSs than Handling-HSs in Agentive contexts, and Adult 3 showed no clear preference. Even more

[29] Visual inspection of the patterns across sessions did not warrant a statistical analysis of the developmental trajectory: for No-Agent events Julio consistently has no strong preference for handshape type, and for Agent events, he consistently strongly prefers Handling-HSs.

Table 1 | For each session except the last, Julio's responses showed a significant association between handshape type and the presence of an Agent in the vignette.

Participant	Age	Number of responses	Pearson Chi-square value ($df = 1$)	p-value two-tailed	Phi coefficient	Effect size
Child	7;4	81	13	0.0003**	−0.40	Medium
Child	7;10–8;5	95	8.8	0.008**	−0.30	Medium
Child	9;11	40	9.72	0.002**	−0.49	Medium
Child	11;4	124	10.92	0.001**	−0.30	Medium
Child	12;8	97	2.59	0.108		
	Total	437				

**$p < 0.01$.*
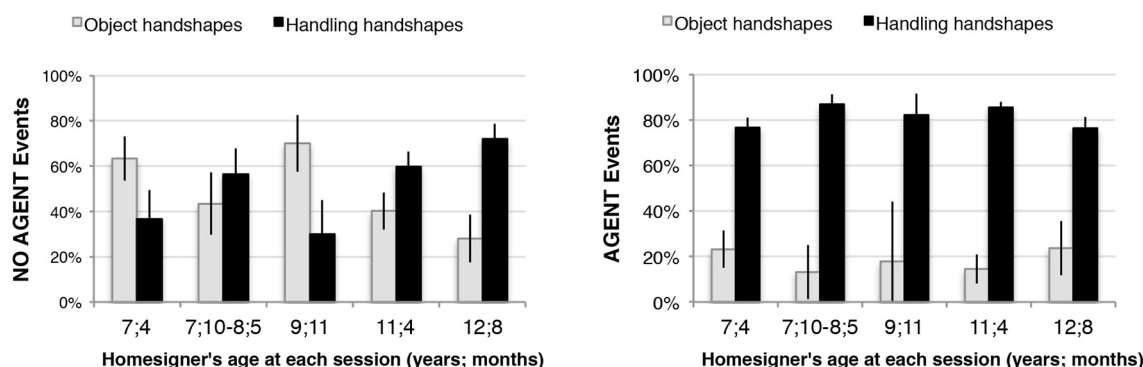
FIGURE 7 | For events with No Agent, Julio produced the expected Object-HSs on average only 49% of the time (gray bars, left chart). However, for events with an Agent he patterned more closely with users of established sign languages, producing Handling-HSs in 82% of items overall (black bars, right chart). These analyses are based on responses to all 10 variations from all 11 objects. Error bars indicate 1 standard error.

Table 2 | All four adult homesigners and the child homesigner in Nicaragua showed a significant association between handshape type (Handling or Object) and vignette type (Agent or Non-Agent), with effect sizes (a measure of strength of association (Cohen, 1988) ranging from medium to large).

Participant	Age	Number of responses	Pearson Chi-square value ($df = 1$)	p-value two-tailed	Phi coefficient	Effect size
Child	7;4–12;8	166	28.84	<0.0001***	−0.42	Medium
Adult 1	20	112	34.53	<0.0001***	−0.56	Large
Adult 2	24	104	6.88	0.01*	−0.26	Medium
Adult 3	29	97	29.67	<0.0001***	−0.55	Large
Adult 4	29	89	34.35	<0.0001***	−0.62	Large

*$p < 0.05$, ***$p < 0.001$.

FIGURE 8 | Comparing Julio with the adult homesigners in Nicaragua in contrastive use of handshape type in events with and without an Agent. For events without an Agent, all four adult homesigners strongly show the established sign language pattern (a preference for Object-HSs, gray bars), whereas Julio does not (left chart). However, in events with an Agent, Julio does strongly show the established sign language pattern (a preference for Handling-HSs, black bars), as does one of the adult homesigners (Adult 4) (right chart). Data from 4 of the 11 objects are included: airplane, book, lollipop, and pen, because these are the objects for which we have comparable data from the adults. Error bars indicate 1 standard error.

striking differences emerge between Julio and the adult homesigners in their responses to Non-Agentive events. Here, while all four adults were more likely to produce Object-HSs than Handling-HSs, Julio did not show a preference for Object-HSs, as do child and adult signers of established and emerging sign languages.

Use of iconic and non-iconic handshape types across groups

We turn now to a comparison of Julio's responses with those of child gesturers and signers, for context (data previously reported in Brentari et al., in press). While the use of iconic handshapes (i.e., Handling- and Object-HSs, vs. other ways of expressing meaning without using speech) might seem obvious to express events with and without agents, especially when explicitly contrasted as they were in these stimuli, this outcome is not inevitable. Brentari et al. (in press) found that while the child signing groups (in Italy and the US) primarily used iconic handshapes in this task (greater than 90%), the child gesture groups in the no-voice condition used them less frequently (71%). These authors also found a developmental shift in this ability: hearing gesturing children in the US produced fewer iconic handshapes than did American adults. Brentari and colleagues also identified a cultural component to the availability of such iconicity:

non-signing adult Italian participants produced iconic handshapes more often than did their American counterparts.

The majority of Julio's event descriptions used iconic handshapes—i.e., Handling, Object or Both. **Figure 9** shows the proportion of iconic and non-iconic *Other* handshapes in Julio's responses compared with those of ASL and LIS child signers, homesigning adults in Nicaragua, and American and Italian child gesturers. The rate of producing an *Other* response (e.g., tracing the path of an object with an index finger) was 6% for Agent events, and 14% for No-Agent events (mean 17%), and is similar to that of the adult homesigners (mean 19%). Signing children produces fewer Other handshapes (3% LIS, 8% ASL), and gesturing children produce more (31% Italian; 25% American). Related analyses using data from these participants is reported in Brentari et al. (in press).

DISCUSSION: USE AND DISTRIBUTION OF ICONIC AND OTHER HANDSHAPES

The use of iconic handshapes in sign languages is pervasive; however, it does not necessarily follow that these iconic uses of handshape are immediately accessible in the manual modality to create a system of linguistic contrasts. As described earlier, there are ways to describe the stimulus events that do not make
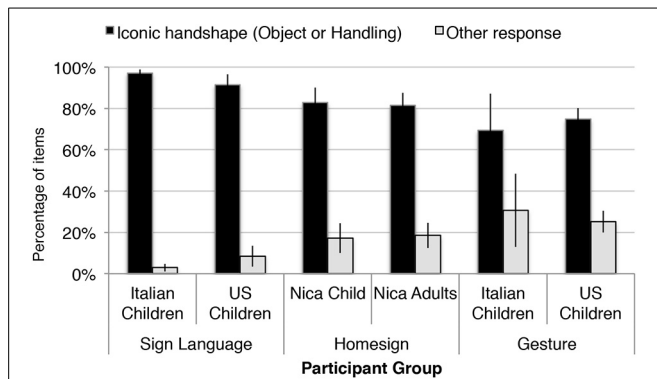
FIGURE 9 | Similarly to the signing Italian and American children, the Nicaraguan homesigning child and adults tend to produce iconic handshapes to describe the events in the stimulus vignettes. This analysis includes all stimulus items involving the objects *airplane* and *lollipop*, and all No-Agent trials for all 11 objects in which there was movement (e.g., the object fell). Error bars indicate 1 standard error.

use of either hand-as-hand or hand-as-object iconicity, such as tracing the shape or trajectory of an object. Thus, the first step in developing a system of handshape oppositions that do morphological work is getting into the ballpark by routinely using iconic handshapes to express such events (Stage 1 of our model). The adult homesigners in Nicaragua have used their homesign systems as their primary language over the course of a lifetime. Julio, the child homesigner, compares favorably with them, as well as with child signers, in his ability to use iconic handshapes when responding to these vignettes. Gesturers use more non-iconic handshapes in their responses to these vignettes, suggesting that using the system as a primary language can trigger the prevalent use of iconic handshape to convey meaning even at a relatively young age. Each individual homesign system displays variability, however, in how fully the morphosyntactic opposition is developed.

In Julio the morphosyntactic opposition is not (yet) fully developed. He showed a distributional difference between the use of Object-HSs and Handling-HSs in all sessions; this contrast emerged despite his strong bias to produce Handling-HSs, and it corresponds to Stage 2 of our model. However, his distribution of handshape types with respect to the presence of an agent differed from that of signers, homesigners, or gesturers. The individual systems of the four adult homesigners in Nicaragua showed a significant association between handshape type (Handling or Object) and vignette type (Agent or Non-Agent), with effect sizes ranging from medium to large. But the contrast in the adult homesigners seems largely driven by their propensity to produce Object-HSs in Non-Agentive events (a pattern similar to that shown by children acquiring ASL), whereas the opposition (such as it is) in the child homesigner is driven by his bias to produce Handling-HSs overall. None of the adult homesigners currently shows this bias toward Handling-HSs, but perhaps they did at an earlier timepoint in the development of their gesture systems. How did it arise in Julio? Is a bias toward Handling handshapes a default starting point for homesigners that is later outgrown? Or, if there is a bias, is the handshape type that is initially preferred

idiosyncratic, with the opposition building from there? Julio's pervasive use of Handling-HSs to describe vignettes both with and without agents sets him apart from the sign and gesture (pantomime) participants. Julio's pattern is not attested in signers (adults or children), who have a language model, nor is it a pattern seen in adult or child gesturers (Brentari et al., in press). He seems to be working out the system using a different strategy than any of the previously studied groups.

In the larger semiotic context of iconicity, Fay et al. (2013, 2014) have proposed that the greater degree of iconicity afforded by the visuo-gestural modality (vs. the auditory-aural modality) allows faster and more efficient development of human communication systems in the absence of language input. This may be true for human communication, broadly construed, but the present results would suggest that while iconicity is clearly available in the visual realm, its use during the creation of a sign language is much more complicated than its wholesale exploitation. As we see here, the general tendency to use iconic handshapes (of any sort) may be a first indication of relationship between handshape and meaning; however, the specific uses of hand-as-object and hand-as-hand iconicity do not get immediately coopted by the system in a sign language-like way. Somewhat counterintuitively, even though the iconicity of this morphological pattern is quite straightforward, and could be achieved simply by imitating the action of the hand as it is engaged in the action, its development is not easy, quick, or obvious.

In summary, children acquiring ASL do not master this distinction until quite late in language development (Schick, 1987; Brentari et al., 2013); similarly, this seems to be a late developing part of the grammar in an emerging language as well, despite its iconic roots. Julio showed a distributional difference between the use of Object-HSs and Handling-HSs in all sessions, but it was a different pattern than we have seen before, one with a strong Handling-HS bias.

STUDY 2: MORPHOPHONOLOGY

In Study 2 we turn to another level of linguistic analysis, morphophonology, and ask whether Julio, during the time period studied, shows evidence of phonological structure in his homesign system. Previous research using this rubric has demonstrated higher complexity handshapes in Object-HSs than in Handling-HSs representations, as is the case for adult and child users of both established (ASL, LIS, CSL-S) and emerging (NSL) languages (Brentari et al., 2012, in preparation). This pattern has also been observed in adult homesigners in Nicaragua (Brentari et al., 2012). The current study has two goals: (1) identifying when in development this distinction emerges by closely examining the handshapes that Julio produced over a 5-year period in response to targeted vignettes and (2) situating this developmental trajectory in the context of the finger group complexity patterns produced by the same four adult homesigners examined in Study 1.

METHODS: MORPHOPHONOLOGY

The participants, stimuli, and procedures were the same as Study 1. In addition to the coding procedures outlined in Study 1, we also transcribed the specific handshapes produced using

the coding system developed by Eccarius and Brentari (2008), as well as the level of complexity of each handshape. This coding system is based on Brentari's (1998) Prosodic Model of Sign Language Phonology, and was developed using handshape forms from 10 different sign languages. Handshape forms were classified according to selected (i.e., active) fingers and joints. To constrain the number of handshape forms, we did not include non-selected (i.e., inactive) fingers in the criterion for a handshape form. For example, a handshape in which the thumb and index finger formed an "O" would be coded as such whether the three non-selected fingers (the middle, ring, and pink fingers) were curled into the palm or left loosely open. We then categorized forms into complexity groups as described in the Introduction.

Low-complexity handshapes received a score of 1, Medium-complexity handshapes a score of 2, and High-complexity handshapes a score of 3. A small number of gesture/sign responses contained more than one finger group and were assigned the score of the highest complexity handshape contained within it. In these cases, one point was added to the complexity score, regardless of how many handshapes were produced (i.e., two distinct finger groups, or more than two). An example of a gesture containing a handshape change that would not count as a change in finger group (but instead reflects a change in joint configuration) is a C-handshape 🖐 that changes to an S-handshape 🤛 (complexity score of 1). An example of a gesture that contains a handshape change that also exhibits different finger groups would be an F-handshape 🤏 that changes to a stacked B handshape 🖐 [(total complexity score of 2): both handshapes are low-complexity, so the baseline score is 1, plus 1 for the change in finger group]. On this metric, complexity scores ranged from 1 to 4, where a score of 4 reflected the highest complexity handshapes (value of 3) plus 1 point in the case of a finger group change.

RESULTS: FINGER GROUP COMPLEXITY
Longitudinal analysis
Figure 10 shows the average finger group complexity for each handshape type (Object, Handling) produced by Julio, the child homesigner, at each session. We first calculated the average finger group complexity for each object (e.g., airplane, book, etc.) within each handshape type, and then averaged across objects. **Table 3** summarizes the results of t-tests comparing the complexity of Object-HSs and Handling-HSs by session. We found that Object-HSs showed higher average finger group complexity than Handling-HSs only for the last two sessions, when Julio was 11;4 and 12;8. This pattern indicates that the established sign language pattern, higher finger group complexity in Object-HSs than Handling-HSs, emerged in the child homesigner prior to the fourth session and persisted into the next session (12;8). The same pattern was found for handshapes that matched the expected morphosyntactic, sign language pattern described in Study 1 and for "violations" of it (i.e., Handling-HSs in No-Agent contexts and Object-HSs in Agent contexts; see Supplementary Material).

One might wonder whether this higher average finger group complexity in Object-HSs was restricted to a small set of objects
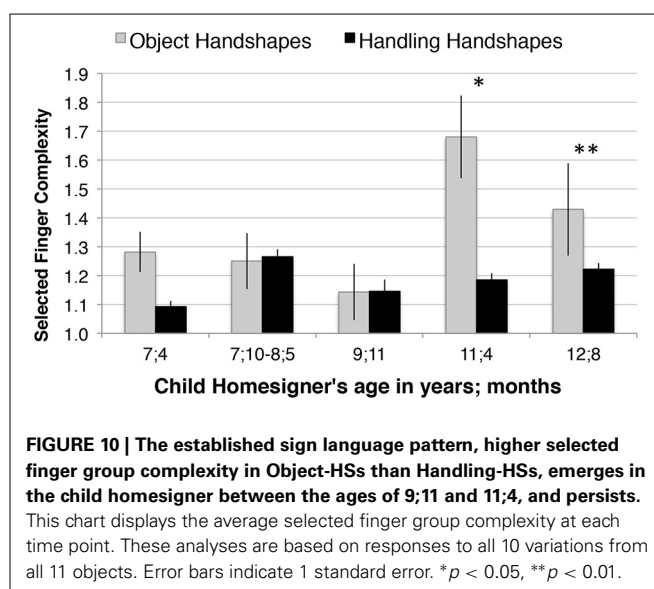

FIGURE 10 | The established sign language pattern, higher selected finger group complexity in Object-HSs than Handling-HSs, emerges in the child homesigner between the ages of 9;11 and 11;4, and persists. This chart displays the average selected finger group complexity at each time point. These analyses are based on responses to all 10 variations from all 11 objects. Error bars indicate 1 standard error. *$p < 0.05$, **$p < 0.01$.

Table 3 | Summary of t-tests comparing the finger group complexity values for Object- and Handling-HSs produced by the child homesigner at each session.

Participant	Age	Number of responses	t-value	df	p-value two-tailed
Child	7;4	81	1.42	9.91	ns
Child	7;10–8;5	94	−0.09	14	ns
Child	9;11	40	−0.22	38	ns
Child	11;4	124	2.44	37.11	0.02*
Child	12;8	97	4.31	95	<0.005**
Child	Total	221			

Non-integer values for degrees of freedom (df) reflect use of the t-test for unequal sample variances. *$p < 0.05$, **$p < 0.01$.

(e.g., airplane, given its relatively complex shape). Julio produced between 15 and 26 handshapes for each of the 11 objects studied. **Figure 11** displays separately the average finger group complexity for the handshapes produced in response to each stimulus object type. For five objects (book, coin, marble, plane, and tweezers), his Object handshapes exhibited higher complexity, and for three objects (cigar, TV, pen) his Handling-HSs showed higher complexity. He produced only simple handshapes (complexity level of 1) for vignettes featuring lollipops, and no Object-HSs for string and tape, thus we were unable to compare Object-HS and Handling-HS complexity for these three objects. Based on the distribution of finger group complexity across these objects, including those that do not appear to demand high complexity handshapes, we conclude that the morphophonological effect observed in the previous analysis is not isolated to specific objects.

Julio compared with adult homesigners
We then compared responses to a subset of four objects (airplane, book, lollipop, and plane) in the child homesigner's last testing session (12;8) ($n = 97$) to those produced by the
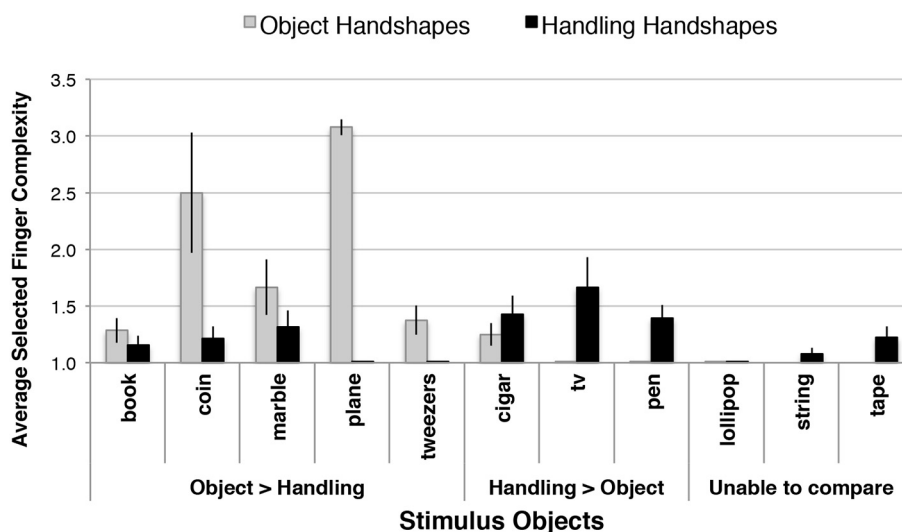
FIGURE 11 | Selected finger group complexity for the handshapes produced in response to each stimulus object type. Julio produced Object handshapes with higher complexity for 5 objects and higher complexity in Handling-HSs for 3 objects. Data from all 11 objects, in all variations, were used in this analysis, but comparisons were unable to be made for 3 objects because Julio did not produce both handshape types for these objects. Error bars indicate 1 standard error.

four adult homesigners previously studied in Nicaragua (total $n = 402$). To standardize the measure of handshape complexity across the different objects so that each object would be weighted equally, we calculated the average finger group complexity across all trials involving each of the four objects described above, and then averaged those, within the sets of Object handshapes and Handling handshapes. Thus, each bar in **Figure 12** reflects the average complexity across four objects[30]. For each adult participant, a t-test for correlated samples was conducted comparing the complexity of Object- and Handling-HSs produced for each object. Adults 1, 2, and 3 produced significantly higher complexity handshapes for Object- than for Handling-HSs, all one-tailed tests: Adult 1 [$t_{(3)} = 3.6$, $p = 0.018$]; Adult 2 [$t_{(3)} = 3.12$, $p = 0.026$]; and Adult 3 [$t_{(3)} = 4.53$, $p = 0.010$]. For Adults 1, 2, and 3, and for the child homesigner, the mean finger group complexity for Object-HSs was greater than or equal to (in two cases) the complexity of the Handling-HSs produced in response to each object (i.e., book, lollipop, pen, and plane). Adult 4 did not show this pattern: [$t_{(3)} = -0.16$, $p = 0.442$]. In accord with the lack of significant difference found in Adult 4, she showed greater complexity in Handling-HSs than Object-HSs for two objects, lollipop and airplane, equal (low) complexity for book, and only produced handshapes with higher complexity in Object-HSs in response to vignettes featuring pens.

Analysis of Specific Handshapes

We now turn to the specific handshapes that were used for each handshape type across participant groups. **Figure 13** shows the

---

[30]The bar for the child homesigner shows handshapes produced in response to three of the four objects, because he did not produce any Object-HSs for vignettes involving pens.
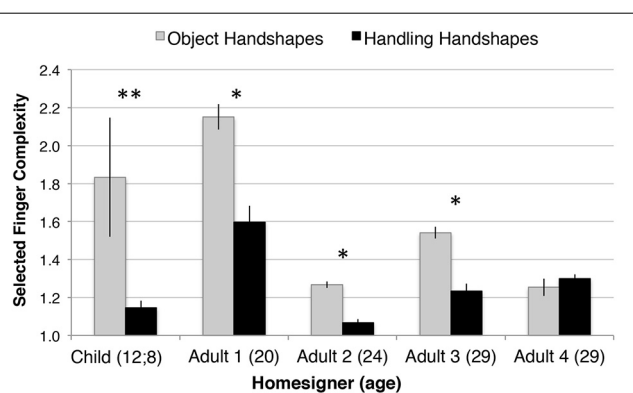
FIGURE 12 | At the last session tested, the child homesigner (age 12;8) showed the established sign language pattern (higher finger group complexity in Object-HSs than Handling-HSs). This pattern was also demonstrated in adulthood by three of the four adult homesigners previously studied in Nicaragua. This analysis includes data from 4 of the 11 study objects: airplane, book, lollipop, and pen, because these are the objects for which we have comparable data from the adult homesigners. Note that a complexity score of 1 is the minimum, and reflects use of the most basic and frequent handshapes observed cross-linguistically. Error bars indicate 1 standard error.*$p < 0.05$, **$p < 0.01$.

handshapes classified as Object- or Handling-HSs that were produced in response to events involving airplanes and lollipops (all 10 variations). The child homesigner's data are from the last testing session, after he had begun to show the sign-like finger group pattern. We also provide the same data from the four adult homesigners in Nicaragua, and for the other groups included in the analysis in Use of iconic and non-iconic handshape types across groups: deaf children acquiring ASL and LIS, and hearing children from the United States and Italy (4–6 years of age)

OBJECT Handshapes used in Event Descriptions involving AIRPLANES

Selected finger group complexity score		Representative Handshapes	Homesigners					Signers		Gesturers	
			Nicaragua					US (ASL)	Italy (LIS)	US (English)	Italy (Italian)
			Adult 1	Adult 2	Adult 3	Adult 4	Child	Children			
1		(handshape)	0.28	0.41	0.67	0.82		0.05	0.27	0.22	0.50
		(handshape)	0.11	0.41	0.11					0.78	0.50
2		(handshape)	0.06								
3		(handshape)		0.18	0.06		0.92		0.73		
		(handshape)	0.39		0.11		0.08	0.95			
		(handshape)			0.06	0.18					
		(handshape)	0.17								
	Total n		13	13	6	2	13	20	11	9	6

OBJECT Handshapes used in Event Descriptions involving LOLLIPOPS

Selected finger group complexity score		Representative Handshapes	Homesigners					Signers		Gesturers	
			Nicaragua					US (ASL)	Italy (LIS)	US (English)	Italy (Italian)
			Adult 1	Adult 2	Adult 3	Adult 4	Child	Children			
1		(handshape)	0.75	0.64	0.79	0.75	1	0.93	0.80	0.60	0.89
		(handshape)		0.29	0.03	0.25		0.07		0.40	0.11
		(handshape)							0.20		
		(handshape)		0.04	0.03						
		(handshape)			0.03						
2		(handshape)		0.04							
		(handshape)			0.03						
3		(handshape)	0.25		0.07						
	Total n		8	28	29	4	1	14	10	10	9

FIGURE 13a | Object Handshapes used in event descriptions involving Airplanes and Lollipops.

who responded using silent gesture. The homesigning and signing groups, but not the gesture groups, produced Object handshapes with Medium- and High-complexity finger groups (**Figure 13a**). Handshapes with Low-complexity finger groups dominate the Handling-HS responses for all groups (**Figure 13b**)[31].

---

[31]Julio's inventory of handshapes in the earlier sessions not shown in these charts (7;4, 7;10–8;5, and 9;11) did not differ from those produced by the hearing, silent-gesturing children.

DISCUSSION: MORPHOPHONOLOGY

We set out to evaluate whether Julio began to show a morphophonological pattern in his use of handshapes during the study period. Specifically, we analyzed Julio's event descriptions for the pattern shown crosslinguistically by adult native signers: higher average finger group complexity in Object-HSs than in Handling-HSs. We found that this morphophonological pattern did emerge in Julio's responses between the ages of 9;11 and 11;4, at which point it was quite robust—it was maintained for

HANDLING Handshapes used in Event Descriptions involving AIRPLANES

Selected finger group complexity score		Representative Handshapes	Homesigners Nicaragua					Signers		Gesturers	
			Adult 1	Adult 2	Adult 3	Adult 4	Child	US (ASL)	Italy (LIS)	US (English)	Italy (Italian)
								Children			
	1		0.22	0.13	0.40	0.42	0.75				
	1						0.25	0.25	0.42	0.60	0.60
	1				0.40						
	1		0.22	0.88	0.20	0.08		0.75	0.50	0.40	0.40
	2		0.56			0.50			0.08		
		Total n	9	8	5	12	9	8	11	5	5

HANDLING Handshapes used in Event Descriptions involving LOLLIPOPS

Selected finger group complexity score		Representative Handshapes	Homesigners Nicaragua					Signers		Gesturers	
			Adult 1	Adult 2	Adult 3	Adult 4	Child	US (ASL)	Italy (LIS)	US (English)	Italy (Italian)
								Children			
	1		0.88	1.00	0.33	0.48	0.72		0.09	0.18	
	1				0.67	0.19	0.11	0.67	0.36	0.64	0.33
	1		0.13				0.17	0.33	0.46		0.50
	2					0.33			0.09	0.18	0.17
		Total n	8	2	3	21	18	6	11	11	6

FIGURE 13b | Handling Handshapes used in event descriptions involving Airplanes and Lollipops.

at least a year (**Figure 10**), and it was not restricted to a small number of object types (**Figure 11**). We also observed higher finger group complexity for Object-HSs than for Handling-HSs in three of the four adult homesigners previously tested in Nicaragua (**Figure 12**). However, the lack of a linguistic model may affect both the timing and strength of the emergence of this pattern—here, we saw the first evidence of it when Julio was 11;4, several years after we observe its emergence in children acquiring ASL. The relationship between phonological and morphological structure is maintained (phonology before morphology), but the timing is delayed, as has been found in other cases of delayed linguistic input (Morford, 2003; Berk and Lillo-Martin, 2012; Ferjan Ramirez et al., 2012).

Both adult and child signers of established sign languages in the US (ASL) and Italy (LIS) showed this phonological pattern. Some sign languages have relatively short histories and are referred to as "emerging sign languages" (Meir et al., 2010). Brentari et al. (in preparation) asked whether the sign-like distribution of finger group complexity requires multiple generations of signers passing down the sign language. They used the approach previously used with signers of established sign languages with child and adult signers of an emerging sign language, Nicaraguan Sign Language (NSL). They found that, indeed, adult signers of NSL showed the established sign language pattern, as did children with 4–6 years of exposure. Thus, this pattern emerges relatively early in development when children are acquiring a sign language from linguistic input[32].

[32] As previously described, we interpret this distribution of finger group complexity as evidence of Julio's developing phonological system. Evidence of phonological development is of course also found in simple lexical items, such as numbers, nouns, and verbs; these uses can and do precede complexity in those same handshapes when they are produced as classifier handshapes (Kantor, 1980). A comparison of the relative timing of these developments in Julio is beyond the scope of this paper; however, we are currently analyzing the handshapes he produces to label objects (comparable to nouns).

None of the hearing adult or child gesturers we have tested, in the US, Italy, or Nicaragua, showed this. Gesturers show very little finger group complexity at all, or else the opposite pattern of higher complexity in Handling-HSs, rather than in Object-HSs (Brentari et al., 2012, in preparation). The high- and medium-complexity finger groups observed in signers' Object-HSs (such as handshapes in which the index and middle fingers are active) were rarely used for Handling-HSs (Eccarius, 2008; Brentari and Eccarius, 2010), even though these handshapes are used in daily life to manipulate certain objects (e.g., holding a baseball or grasping a small teacup by its handle).

In summary, these results, combined with the present results from Julio, suggest that the morphophonological pattern does not appear to require linguistic input in order to emerge, and that it is not inevitable when using the manual modality.

OVERALL DISCUSSION

We began by identifying universals in the ways that sign languages use two aspects of handshape: (1) handshape type and (2) finger group complexity, to mark linguistic contrasts. Importantly, in sign languages handshape type—the systematic use of Object-HSs vs. Handling-HSs, depending on the presence of an Agent—is grammatical: the distribution of handshape types is associated with a meaning contrast (Agentivity) and thus constitutes a morphosyntactic system. These handshape types are also systematically associated with contrasting levels of average finger group complexity—higher in Object-HSs than in Handling-HSs—and are in this sense morpho-phonological. Recall that the phonological patterns we observed are embedded within the morphosyntactic structure of the classifier constructions of sign languages. We discovered a degree of convergence between these language universals for sign languages that have classifier systems, and in the behavior of individual child and adult homesigners. We then attempted to identify the source of this convergence—specifically, whether it could be attributed to shared constraints on iconicity. We turn now to an integrative discussion of how the distributions of handshape type (Object vs. Handling-HSs) and finger group complexity that we observed in the child homesigner Julio constitute linguistic, and specifically, morphophonological development, rather than mere elaborations of iconic patterns that emerge from the affordances of how humans interact with objects. Specifically, we show how the results and analyses from Studies 1 and 2 serve to evaluate our proposed model of the emergence of systematic distributions of handshape type and finger group complexity in the absence of conventional linguistic input.

Returning to the *Distributional Model* proposed in the Introduction, we can now insert our findings from the two studies presented:

Stage 1: Recognizing Handshape Type as an aspect of handshape form that can be utilized for grammatical purposes.

Finding: Julio predominantly uses Object and Handling (iconic) handshapes, rather than neutral handshapes or full-body expressions, to describe events involving objects and manipulation (see **Figure 9**). Rooted in *hand-as-object* and *hand-as-hand* iconicity, they can be thought of as the raw materials from which a morphological system can be constructed.

Stage 2: Distinguishing the distribution of Object and Handling handshapes in one's system; associating one handshape type with one event type and the other handshape type to the other event type *to some degree*. This association does not have to be complete, nor does it have to be present for both handshape types/event types, in order for a contrast to emerge between these handshape classes.

Finding: Julio makes this distinction, manifested by his different distributions of handshape type to express events with and without an Agent (see **Table 1** and **Figure 7**). We interpret his association of Handling handshapes with Agentive events, and to a lesser degree his association of Object handshapes with Non-agentive events as evidence of this. These two handshape classes thus lay the foundation for the phonological pattern to appear.

Stage 3: Organizing phonological properties with regard to handshape classes. This organization need not be independent from the morphological category to be phonological (see **Figure 3**). Note that a contrast between agentive and non-agentive events that is encoded in handshape is all that is needed before this **morphophonological** pattern can develop.

Finding: By the second-to-last session, Julio's Object-HSs displayed higher average finger group complexity than his Handling-HSs, and this pattern was not restricted to a small number of objects (see **Table 3** and **Figures 10**, **11**).

Stage 4: Using Handshape Type systematically and oppositionally to mark the presence or absence of an agent, as is observed in the **morphosyntax** of classifier constructions in established sign languages[33].

Finding: In contrast to the robustness of Julio's homesign system with respect to Stages 1 through 3 of the model, he does not show evidence of using handshape to mark this grammatical opposition (see **Table 1** and **Figure 7**); nor do two of the four adult homesigners whose handshapes have been studied using the same stimuli and analytic procedures (see **Table 2** and **Figure 8**).

Interestingly, the adult homesigners show a range of outcomes with respect to the model. Like Julio at the end of the study period, Adults 2 and 3 have reached Stage 3 (showing the morphophonological pattern, but not the full morphosyntactic opposition). Adult 1 has reached Stage 4 (showing the morphophonological and the full morphosyntactic opposition found in sign languages). Adult 4's performance is atypical: though she shows the morphosyntactic opposition of Stage 4, she did not develop the morphophonological pattern (Stage 3). One explanation might be that she did not develop a sufficiently large or complex inventory of handshapes that could then be used differentially in Object and Handling handshapes. We leave this for future work.

---

[33]The theoretical model for this morphosyntactic contrast is laid out in Benedicto and Brentari (2004). In practice, we have used the performance of native-signing adult signers as criteria for this contrastive use of handshape type in the grammar of established sign languages; while adult signers demonstrate a clear contrast in their use of handshape type, they do not uniformly produce the expected forms 100% of the time (Brentari et al., 2012), underscoring the importance of obtaining behavioral data that converges with theoretical predictions.

Within this productive subsystem of a sign language, the morphophonological pattern develops earlier than the morphosyntactic one in child signers as well as in this single Nicaraguan child homesigner whose exposure to NSL was extremely sparse and sporadic. The results of the two studies, and their relation to the proposed model, are summarized in **Table 4**.

One can ask whether Julio would behave more like a child signer in his development (phonology before morphosyntax) or more like some adult gesturers who produce the target handshape distinction based on the presence of an agent without a corresponding phonological level of structure. Julio apparently behaves more like a child signer than an astute Italian gesturer: he produces the phonological pattern in finger group, but not the expected pattern in handshape type. An important point regarding comparing homesigners and gesturers on the same task is that homesigners come to the task with experience of using their system on a daily basis to express a variety of meanings and grammatical contrasts, while the gesturers are inventing their responses on the spot. The homesigners, therefore, are constantly balancing and integrating multiple aspects of their systems, while the gesturers are presented with a single, specific communicative task that has been tailored into a bite-sized chunk. Differences between Julio and adult Italian gesturers exemplify the consequences of moving beyond this restricted domain of solving a single communication problem by expressing Agentive vs. Non-Agentive events, to the complexity of trying to solve the multi-dimensional problem of expressing a variety of contrasts simultaneously in the creation of a linguistically organized system. The latter is the task faced by a single homesigner in the absence of linguistic input.

PHONOLOGICAL DEVELOPMENT AND MORPHOLOGY IN SIGNED AND SPOKEN LANGUAGES

Previous studies of the acquisition of phonology and handshape in sign languages have focused almost exclusively on the timing and patterns of acquisition of lexical nouns and verbs, or on the acquisition of semantic classifier forms, such as Object/Entity classifiers. However, most studies have converged on a common conclusion, that different formational parameters tend to be acquired in a piecemeal fashion (i.e., different timing for the acquisition of location and handshape configuration; see, for example, Boyes Braem, 1990; Marentette and Mayberry, 2000; Meier, 2006; Ortega and Morgan, 2010). Further, the core lexicon is not the only place within a grammatical system that one might look for evidence of phonological structure (as in, for example, Kantor, 1980; Fish et al., 2003; Eccarius, 2008)—our study uniquely addresses phonological patterns that take into account the morphosyntactic function of the classifier construction, and exist beyond the domain of the lexicon. Considering phonological structure more broadly, then, our observations of Julio indicate that phonology appears relatively early, in the form of contrastively used finger group complexity.

The relatively late acquisition of morphosyntactic patterns in ASL and other sign languages, and the close interplay between

Table 4 | Brief descriptions of the stages of the Distributional Model and summary of results from Studies 1 and 2.

		Users of established sign languages	Homesigning participants					
			Child 7;4	Child 11;4	Adult 1 20 years	Adult 2 24 years	Adult 3 29 years	Adult 4 29 years
Predominant use of iconic handshape types (Object- and Handling-HSs) (Stage 1)		Yes	Yes	Yes	Yes	Yes	Yes	Yes
Morpho-syntax	Association between handshape and agentivity (Stage 2)	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	Agentive events	Handling-HSs	Handling-HSs	Handling-HSs	Handling-HSs	Object-HSs	No strong preference	Handling-HSs
	Non-agentive events	Object-HSs	No strong preference	No strong preference	Object-HSs	Object-HSs	Object-HSs	Object-HSs
Morpho-phonology	Greater complexity for Object-HSs than Handling-HSs (Stage 3)	Yes	No	Yes	Yes	Yes	Yes	No
Systematic opposition of handshape type between agentive and non-agentive events (Stage 4)		Yes	No	No	Yes	No	No	Yes

phonology and morphology, is not unique to sign languages, but is also found in spoken languages (e.g., MacWhinney, 1978; Levinger-Gottlieb, 2007; Ravid and Schiff, 2009). Since previous work demonstrated that three of four homesigners tested in adulthood already showed the morphophonological pattern that Julio displayed here in later sessions, we can also now put this finding into the broader context of phonological development in the absence of a linguistic model. The present work represents the first study of phonological and morphosyntactic development in the use of handshape over time in a homesigner of any age who has yet to be immersed in a sign language environment. It also adds to a very small literature addressing the development over time of any linguistic structure in homesign systems that continue to be used as primary languages beyond early childhood[34].

ICONICITY AND MORPHOSYNTAX

If hand-as-hand and hand-as-object types of iconicity are widely accessible to all populations, we might expect anyone who responded to these vignettes to show the morphosyntactic-*like* opposition described in Stage 4 100% of the time. Indeed, the participants across all language groups and ages in these studies could have achieved the sign-like morphosyntactic pattern by simply mimicking the actions of the human agent in the Agentive events. Likewise, they could have succeeded in the No-Agent events by refraining from inserting an Agent into the event, that is, by using any non-Handling handshape to express the arrangement or movement of the object(s) in the vignettes (e.g., a simple handshape such as, ☞). But this is not the pattern we have observed. Julio failed to exploit fully and equally the iconicity present in Handling- and Object-HSs. One possibility is that he did not grasp the distinction between Agentive vs. Non-Agentive events. While we cannot rule this out, Julio certainly appeared to understand the task; the instructions given to all participant groups are quite minimal, namely "describe what you see." His strikingly different handshape distributions across the two types of events suggests that he was sensitive to the presence of an agent[35]. Moreover, we have seen this pattern across a number of other populations: in the adult homesigners in Nicaragua

also reported here; among Cohort 1 and 2 signers of NSL and native signers of ASL (Goldin-Meadow et al., under review), as well as among children acquiring ASL and LIS (Brentari et al., in press).

Brentari et al. (in press) argue that while cognitive and cultural factors influence the use of the handshape for this purpose, there is also a strong linguistic component to this opposition, as argued in Benedicto and Brentari (2004); thus handshape used in a systematic, motivated way to mark agency and transitivity, as described above, is *both* a specifically motivated iconic pattern (one that even some gesturers can discern), but used in the service of grammar in sign languages (cf. Meir et al., 2013).

The argument that Julio only "overuses" hand-as-hand iconicity, on our view, constitutes evidence for different subtypes of iconicity and is the foundation for our claim that Julio's system is moving beyond what is offered by perceptual/cognitive affordances and into a linguistic realm. This linguistification is driven by the continued need to develop the system itself (in the absence of a linguistic model), and the consequent requirement that forms exist in relationship to other forms, rather than just being associated with meanings in the world. Converging evidence comes from other studies of homesign systems developed by children in the US and China, which exhibit morphological structure, i.e., handshape categories and motion categories that combine productively to create new signs (e.g., Goldin-Meadow et al., 1995, 2007). It is more appropriate to characterize Julio's bias toward Handling-HSs as an incomplete re-organization of iconicity by his grammar. Julio uses as many iconic handshapes as signers do at the same age. Moreover, the predicted sign language pattern is based on iconicity as well (the two types: hand-as-hand and hand-as-object), and Julio shows a different distribution of his use of iconicity.

ICONICITY AND MORPHOPHONOLOGY

Like the morphosyntactic pattern, the morphophonological pattern is also iconic, though in a different, more indirect way, and with a less direct relationship between the stimuli that the participants saw and the task they were asked to perform. If there were an obvious solution to expressing that iconicity, presumably gesturers (certainly adults) would also demonstrate it, but the evidence from gesturers in three different cultures suggests that they do not (Brentari et al., 2012, in preparation). Our study specifically investigated the use of handshape to mark this opposition. However, it is possible to imagine other parameters of sign formation, such as movement, representing the first step toward marking the agentive/non-agentive distinction. One form this might take would be to use "contact" movements (i.e., movements with a final "bump" that highlight the "surface" on which an object rests) to express stative events vs. events with movement, which would overlap some with the agentive/agentive distinction, but not completely (e.g. objects that fall). We leave this for future work.

Several researchers have argued for a crucial role of iconicity in the development of structure in the manual modality, in both emerging systems such as homesign as well as in established sign languages (see, for example, Cuxac's (1999) work on iconicity from a semiotic perspective in French Sign Language).

---

[34]Though see Goldin-Meadow et al. (1994) regarding a noun-verb distinction in child homesigners, Goldin-Meadow and Mylander (1990) on gesture order and other patterns in child homesigners from 1;4 to 5;9; Morford (2003) for a longitudinal study of the use of handshape to express motion events in two adolescent homesigners who were recently immersed in ASL. Berk and Lillo-Martin (2012) analyzed the two-word utterances produced over time by two deaf children who began acquiring ASL in an immersion context at the ages of 5;9 and 6;0.

[35]One alternative hypothesis is that the salience of human agents accounts for Julio's use of Handling-HSs. Julio may have inferred that someone put the objects in the frame though the agent was not observed doing so. While we cannot definitively rule this out, neither signers nor hearing individuals gesturing silently in the US and Italy did this; in fact, they used Object-HSs more often for Agentive vignettes than the reverse pattern exhibited by Julio. A second alternative explanation is his young age. Kaplan (1968) found that children asked to show how they would use non-present objects did not robustly use handling gestures until they were 12 years old, suggesting that Julio's early use of Handling-HSs in Agentive contexts does not reflect "default" imitation or an immature pattern.

It has been argued that users of homesign must rely on iconicity in order to maintain transparency and comprehensibility[36] (e.g., Goldin-Meadow, 2003; Fusellier-Souza, 2006). In the domain of experimental semiotics, Fay et al. (2013, 2014) have argued that iconicity can facilitate human communication in the absence of linguistic input. This use of iconicity is pre-Stage 1 in in our model, because different ad hoc strategies for employing different kinds of iconicity might suffice for a single, relatively constrained communicative task, but not for a primary language that has to serve many functions. And while iconicity facilitates communication, the present results suggest that iconicity is not sufficient to build a linguistic system.

LANGUAGE EMERGENCE AND EVOLUTION

The researchers studying another recently emerged language, Al-Sayyid Bedouin Sign Language (ABSL) have observed that morphology appears to be developing more quickly than phonological structure (e.g., Sandler et al., 2005; Padden et al., 2010; Sandler et al., 2011). The morphological structure they have studied relates to verb (person) agreement, which appears in signers of the first generation of ABSL. The present results from Julio, a child homesigner studied longitudinally, seem to indicate the opposite: that phonological structure can appear relatively early, while morphological development requires more time. We can think of at least three ways to reconcile these apparently contradictory findings: (1) Morphology and phonology are expressed in different ways in different subsystems of the grammar. Our study focused on the use of handshape to mark contrasts in agentivity, while the morphological aspects of verb/person agreement involve the movement of verb signs in signing space. Further, examinations of phonology in the context of morphology, comparable to those under study here, have not been reported in ABSL. (2) Perhaps the type of structure they put forward as morphological, reflecting the notion of "body-as-subject" and involving movements of verbs along the midsagittal plane of the body, are anaphoric at the level of discourse rather than at the morphosyntactic level. These authors also suggest this as a possible explanation of their findings. (3) ABSL is a village sign language, in which deaf and hearing users interact with each other, and many deaf individuals use the language as their primary language. This is a different sociolinguistic setting from that of homesign, the focus of the current study. Julio and the adult homesigners in Nicaragua do not experience the same pressures to conventionalize their linguistic systems with the hearing people around them, who do not use the system as a primary language. Without this pressure, perhaps phonological complexity has a higher probability of emerging. It may be that regular interactions with individuals who do not use the manual system as their primary language (i.e., hearing communication partners) hinder the homesigner's internal consistency (see Richie et al., 2014; Goldin-Meadow et al., under review, on the

conventionalization of lexical items in the individual homesign family groups).

Advantages of our comparative approach include the ability to separately identify the contributions of several factors to the emergence of linguistic structure: linguistic input; stage of development; use of the manual modality as a primary language for homesigners and signers vs. a one-time occurrence for hearing gesturers; and culture. To address the role of repeated and habitual use of the manual modality to express ideas/concepts that would typically be expressed in speech in hearing individuals, in ongoing work we are following hearing, non-signing gesturers who do not have regular contact with homesigners, as well as homesigners' regular communication partners over time.

Homesign systems differ in significant ways from sign languages used by a community of Deaf people (e.g., Spaepen et al., 2011, 2013 on the lack of a count list; and Richie et al., 2014 and Goldin-Meadow et al., under review on the slower conventionalization of lexical items). However, the available evidence strongly indicates that homesign more closely resembles sign language than it does gesture. This evidence, unsurprisingly, comes from studies of linguistic structures that do not involve conventionalization among a community of users (e.g., Coppola and Senghas, 2010 on pronouns; Brentari et al., 2012 on morphophonology; Coppola et al., 2013 on plurals). In accord with the findings summarized above, careful consideration of the relationships among the individual structural components exhibited in homesign reveal that constraints across different levels of linguistic analysis are much weaker than they are in either emerging or established languages, which have the benefit of a linguistic community and/or linguistic input. When linguistic input is available, it apparently constrains multiple levels of linguistic structure simultaneously, but without linguistic input, cohesion and integration across components of the grammar is less apparent, and we can see the piecemeal development of sub-parts of both morpho-syntax and morpho-phonology.

CONCLUSION

Perhaps because of the affordances offered by a language system using the manual modality, the notion of iconicity as a highly complex, multi-layered set of phenomena that are utilized in distinctly different ways in sign languages is often not fully appreciated. Brentari (2007) notes that while it is clear that iconic sources can be identified for many aspects of sign language structure, it is also evident that "arbitrary formal structure is present and observable at every level of SL grammar." Indeed, this notion is echoed in the present findings from one child homesigner, in which we observe a less iconic form (contrastive use of finger group complexity) emerging earlier in development than a type of iconicity that appears more straightforward (namely, hand-as-hand iconicity). We interpret his lack of exploiting the hand-as-object iconicity as a consequence of the fact that, as a homesigner, he is building a grammatical system from non-linguistic gestural input.

These findings constitute evidence that individual components of a phonological system can exist before there is a full phonological system (e.g., one that includes minimal pairs and assimilation rules). In this regard, these results accord with previous

---

[36] Though also see anecdotal examples such as the following, from Feldman (1975) (and others discussed by Meier, 1982), in which a child homesigner developed a gesture referring to ice cream that used licking as its iconic base, and continued to use that form even in situations in which the ice cream was in a bowl, and no licking was involved.

findings with adult homesigners (Brentari et al., 2012) and children acquiring an established sign language (Brentari et al., 2013). Some adult gesturers achieve Stage 1 of our Distributional Model, where handshape is meaningfully manipulated (Brentari et al., in press) and signing children and adult homesigners achieve all four stages of the model, where the opposition of handshape in phonology and morphology is clear (Brentari et al., 2012, 2013, under review). The evidence we have described here shows that a single child homesigner has achieved Stage 3 (morpho)-phonology), but not Stage 4 (morphosyntax). The particular manifestation of these components in both child and adult homesign systems, while not deterministic, nevertheless generally accords with the associations seen in emerging and established sign languages. In other words, while a homesigner will not achieve the same level of linguistic sophistication, in both homesign and sign languages, iconicity is dismantled and reassembled in the service of a multi-componential system. Further, these distributional patterns do not reflect the patterns observed in the gestures produced by hearing people to describe these vignettes in the manual modality, who may astutely exploit available patterns of iconicity and their life experience with co-speech gesture and a spoken language and apply these skills to a specific gestural "problem" presented in a controlled task.

While it is difficult to generalize from a set of five case studies, we take the child and adult homesign findings as an existence proof that some aspects of morphophonology and morphosyntax can develop within an individual who is not acquiring a conventional language. Considered in conjunction with related studies, these findings also suggest that iconicity, in the sense of the hand representing the hand of an agent, or representing an object's movement or properties, does not entirely drive these linguistic developments. Nor is this iconicity easily accessible to individuals gesturing without voice (pantomime) who do not routinely communicate in this fashion. Ongoing analyses of the gesture descriptions produced by the communication partners of the child and adult homesigners offer an opportunity to distinguish these factors. Taken together, this body of work suggests that, while handling and object handshapes are ubiquitous and iconic, the various cognitive and linguistic roles these handshapes can assume cannot be conflated and must be investigated independently, and more importantly, analyzed distributionally.

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://www.frontiersin.org/journal/10.3389/fpsyg.2014.00830/abstract

REFERENCES

Aronoff, M., Meir, I., Padden, C., and Sandler, W. (2003). "Classifier constructions and morphology in two sign languages," in *Perspectives on Classifier Constructions in Sign Languages*, ed K. Emmorey (Mahwah, NJ: Lawrence Erlbaum), 53–84.

Benedicto, E., and Brentari, D. (2004). Where did all the arguments go?: Argument-changing properties of classifiers in ASL. *Nat. Lang. Linguist. Theory* 22, 743–810. doi: 10.1007/s11049-003-4698-2

Berk, S., and Lillo-Martin, D. (2012). The two-word stage: motivated by linguistic or cognitive constraints? *Cogn. Psychol.* 65, 118–140. doi: 10.1016/j.cogpsych.2012.02.002

Bernardino, E. L. A. (2006). What Do Deaf Children Do When Classifiers Are Not Available? *The Acquisition of Classifiers in Verbs of Motion and Verbs of Location in Brazilian Sign Language (LSB)*. Ph.D. Dissertation, Boston University.

Borer, H. (2005). *Structuring Sense: Volume II: The Normal Course of Events.* Oxford: Oxford University Press.

Boyatzis, C., and Watson, M. (1993). Children's symbolic representation of objects through gestures. *Child Dev.* 64, 729–735. doi: 10.2307/1131214

Boyes Braem, P. (1990). "Acquisition of the handshape in American sign language," in *From Gesture to Language in Hearing and Deaf Children*, eds V. Volterra and C. J. Erting (New York, NY: Springer Verlag), 107–127.

Brentari, D. (1998). *A Prosodic Model of Sign Language Phonology.* Cambridge, MA: MIT Press.

Brentari, D. (2007). "Sign language phonology: issues of iconicity and universality," in *Verbal and Signed Languages*, eds E. Pizzuto and R. Simone (Berlin: Mouton de Gruyter), 59–80.

Brentari, D. (2011). "Handshape in sign language phonology," in *The Blackwell Companion to Phonology*, eds M. van Oostendorp, C. Ewen, E. Hume, and K. Rice (Oxford: Blackwell Publishing Ltd), 195–222.

Brentari, D., and Coppola, M. (2012). *What Sign Language Creation Teaches us About Language.* WIREs Cognitive Science.

Brentari, D., Coppola, M., Jung, A., and Goldin-Meadow, S. (2013). Acquiring word class distinctions in American Sign Language: evidence from handshape. *Lang. Learn. Dev.* 9, 130–150. doi: 10.1080/15475441.2012.679540

Brentari, D., Coppola, M., Mazzoni, L., and Goldin-Meadow, S. (2012). When does a system become phonological? Handshape production in gesturers, signers, and homesigners. *Nat. Lang. Linguist. Theory* 30, 1–31. doi: 10.1007/s11049-011-9145-1

Brentari, D., Di Renzo, A., Keane, J., and Volterra, V. (in press). Cognitive, cultural, and linguistic sources of a handshape distinction expressing agentivity. *Topics Cogn. Sci.*

Brentari, D., and Eccarius, P. (2010). "Handshape contrasts in sign language phonology," in *Sign Languages: a Cambridge Language Survey*, ed D. Brentari (Cambridge: Cambridge University Press), 284–311.

Brentari, D., Tang, G., and Benedicto, E. (2001). "A comparative study of agentive and non-agentive use of classifiers in HKSL and ASL," in *Paper Presented at the Conference on Asian Sign Languages, Deaf Culture and Deaf Education* (Hong Kong: Chinese University of Hong Kong).

Carrigan, E., and Coppola, M. (2012). "Mothers do not drive structure in adult homesign systems: evidence from comprehension," in *Proceedings of the 34th Annual Conference of the Cognitive Science Society*, eds N. Miyake, D. Peebles, and R. P. Cooper (Sapporo: Cognitive Science Society), 1398–1403. Available online at: http://mindmodeling.org/cogsci2012/papers/0249/paper0249.pdf

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences. 2nd Edn.* New York, NY: Academic Press.

Coppola, M. (2014). *From Iconic Handshapes to Grammatical Contrasts: Video Examples for Coppola and Brentari Frontiers Article.* Figshare.

Coppola, M., and Newport, E. L. (2005). Grammatical subjects in home sign: abstract linguistic structure in adult primary gesture systems without linguistic input. *Proc. Natl. Acad. Sci. U.S.A.* 102, 19249–19253. doi: 10.1073/pnas.0509306102

Coppola, M., and Senghas, A. (2010). "Deixis in an emerging sign language," in *Sign Languages: A Cambridge Language Survey*, ed D. Brentari (Cambridge: Cambridge University Press), 543–569.

Coppola, M., Spaepen, E., and Goldin-Meadow, S. (2013). Communicating about quantity without a language model: number devices in homesign grammar. *Cogn. Psychol.* 67, 1–25. doi: 10.1016/j.cogpsych.2013.05.003

Crasborn, O., and Sloetjes, H. (2008). "Enhanced ELAN functionality for sign language corpora," in *Proceedings of LREC 2008, Sixth International Conference on Language Resources and Evaluation* (Marrakech).

Cuxac, C. (1999). "The expression of spatial relations and the spatialization of semantic relations in French Sign Language," in *Language Diversity and Cognitive Representations*, eds C. Fuchs and S. Robert (Amsterdam: John Benjamins), 123–142.

Demey, E., and van der Kooij, E. (2008). Phonological patterns in a dependency model: allophonic relations grounded in phonetic and iconic motivation. *Lingua* 118, 1109–1138. doi: 10.1016/j.lingua.2007.12.003

Eccarius, P. (2008). *A Constraint-Based Account of Handshape Contrast in Sign Languages.* Ph.D. Dissertation, Purdue University, West Lafayette, IN.

Eccarius, P., and Brentari, D. (2008). Handshape coding made easier: a theoretically based notation for phonological transcription. *Sign Lang. Linguist.* 11, 69–101. doi: 10.1075/sl&U38;l.11.1.11ecc

ELAN. *EUDICO Linguistic Annotator.* Available online at: http://tla.mpi.nl/tools/tla-tools/elan/

Emmorey, K. (ed.). (2003). *Perspectives on Classifier Constructions in Sign Languages.* Mahwah, NJ: Lawrence Erlbaum.

Emmorey, K., and Herzig, M. (2003). "Categorical vs. gradient properties in classifier constructions in ASL," in *Perspectives on Classifier Constructions in Sign Languages*, ed K. Emmorey (Mahwah, NJ: Lawrence Erlbaum Associates), 221–246.

Fay, N., Arbib, M., and Garrod, S. (2013). How to bootstrap a human communication system. *Cogn. Sci.* 37, 1356–1367. doi: 10.1111/cogs.12048

Fay, N., Lister, C. J., Ellison, T. M., and Goldin-Meadow, S. (2014). Creating a communication system from scratch: gesture beats vocalization hands down. *Front. Psychol.* 5:354. doi: 10.3389/fpsyg.2014.00354

Feldman, H. M. (1975). *The Development of a Lexicon by Deaf Children of Hearing Parents: or, There's More to Language than Meets the Ear.* Doctoral Dissertation, University of Pennsylvania, Philadelphia, PA.

Ferjan Ramirez, N., Lieberman, A. M., and Mayberry, R. I. (2012). The initial stages of language acquisition begun in adolescence: when late looks early. *J. Child Lang.* 40, 1–24. doi: 10.1017/S0305000911000535

Fish, S., Morén, B., Hoffmeister, R., and Schick, B. (2003). "The acquisition of classifier phonology in ASL by deaf children," in *Proceedings of the 27th annual Boston University Conference on Language Development of (BUCLD 27)*, Vol. 1 (Somerville, MA: Cascadilla Press), 252–263.

Fusellier-Souza, I. (2006). Emergence and development of sign languages: from a semiogenetic point of view. *Sign Lang. Stud.* 7, 30–56. doi: 10.1353/sls.2006.0030

Goldin-Meadow, S. (2003). *The Resilience of Language: What Gesture Creation in Deaf Children Can Tell us About How All Children Learn Language.* New York, NY: Psychology Press.

Goldin-Meadow, S., Butcher, C., Mylander, C., and Dodge, M. (1994). Nouns and verbs in a self-styled gesture system: what's in a name? *Cogn. Psychol.* 27, 259–319. doi: 10.1006/cogp.1994.1018

Goldin-Meadow, S., Mylander, C., and Butcher, C. (1995). The resilience of combinatorial structure at the word level: morphology in a self-styled gesture system. *Cognition* 56, 195–262. doi: 10.1016/0010-0277(95)00662-I

Goldin-Meadow, S., and Mylander, C. (1990). Beyond the input given: the child's role in the acquisition of language. *Language* 66, 323–355. doi: 10.2307/414890

Goldin-Meadow, S., Mylander, C., and Franklin, A. (2007). How children make language out of gesture: morphological structure in gesture systems developed by American and Chinese deaf children. *Cogn. Psychol.* 55, 87–135. doi: 10.1016/j.cogpsych.2006.08.001

Hara, D. (2003). *A Complexity-Based Approach to the Syllable Formation in Sign Language.* Doctoral Dissertation, University of Chicago, Chicago, IL.

Janis, W. (1992). *Morphosyntax of the ASL Verb Phrase.* Unpublished Ph.D. Dissertation, State University of New York, Buffalo, NY.

Kaplan, E. (1968). *Gestural Representation of Implement Usage: an Organismic Developmental Study.* Doctoral Dissertation, Clark University, Worcester, MA.

Kantor, R. (1980). The acquisition of classifiers in American Sign Language. *Sign Lang. Stud.* 28, 193–208. doi: 10.1353/sls.1980.0000

Kegl, J. (1990). "Predicate argument structure and verb-class organization in the ASL Lexicon," in *Sign Language Research: Theoretical Issues*, ed C. Lucas (Washington, DC: Gallaudet University Press), 149–175.

Klima, E. S., and Bellugi, U. (1979). *The Signs of Language.* Cambridge, MA: Harvard University Press.

Levinger-Gottlieb, M. (2007). *The Acquisition of Hebrew Plurals: A Morpho-Phonological Account.* Masters Thesis, Tel-Aviv University, Tel-Aviv.

Marentette, P. F., and Mayberry, R. I. (2000). "Principles for an emerging phonological system: a case study of early American Sign Language acquisition," in *Language Acquisition by Eye*, eds C. Chamberlain, J. Morford, and R. Mayberry (Mahwah, NJ: Lawrence Erlbaum Associates), 71–90.

MacWhinney, B. (1978). *The Acquisition of Morphophonology (Vol. 43, Monographs for the Society for Research in Child Development).* Oxford: Blackwells.

Mazzoni, L. (2009). *Classificatori e Impersonamento nella Lingua dei Segni Italiana.* Pisa: Plus.

Meier, R. P. (1982). *Icons, Analogues, and Morphemes: The Acquisition of Verb Agreement in ASL.* Doctoral Dissertation, University of California, San Diego, CA.

Meier, R. P. (2006). "The form of early signs: explaining signing children's articulatory development," in *Advances in the Sign Language Development of Deaf Children*, eds B. Schick, M. Marschark, and P. E. Spencer (New York, NY: Oxford University Press), 202–230.

Meir, I. (2010). Iconicity and metaphor: constraints on metaphorical extension of iconic forms. *Language* 86, 865–896. Available online at: http://muse.jhu.edu/journals/lan/summary/v086/86.4.meir.html

Meir, I., Padden, C., Aronoff, M., and Sandler, W. (2007). The body as subject. *J. Linguist.* 43, 531–563. doi: 10.1017/S0022226707004768

Meir, I., Padden, C., Aronoff, M., and Sandler, W. (2013). Competing iconicities in the structure of languages. *Cogn. Linguist.* 24, 309–343. doi: 10.1515/cog-2013-0010

Meir, I., and Sandler, W. (2007). *Language in Space: The Story of Israeli Sign Language.* New York, NY: Lawrence Erlbaum Associates.

Meir, I., Sandler, W., Padden, C., and Aronoff, M. (2010). "Emerging sign languages," in *Oxford Handbook of Deaf Studies, Language, and Education*, Vol. 2, eds M. Marschark and P. Spencer (Oxford: Oxford University Press), 267–280.

Morford, J. P. (2003). Grammatical development in adolescent first language learners. *Linguistics* 41, 681–721. doi: 10.1515/ling.2003.022

Namy, L., Campbell, A. L., and Tomasello, M. (2004). The changing role of iconicity in non-verbal symbol learning: A U-shaped trajectory in the acquisition of arbitrary gestures. *J. Cogn. Dev.* 5, 37–57. doi: 10.1207/s15327647jcd0501_3

O'Reilly, A. W. (1995). Using representations: comprehension and production of action with imagined objects. *Child Dev.* 66, 999–1010.

Ormel, E., Hermans, D., Knoors, H., and Verhoeven, L. (2009). The role of sign phonology and iconicity during sign processing: the case of deaf children. *J. Deaf Stud. Deaf Educ.* 14, 436–448. doi: 10.1093/deafed/enp021

Ortega, G., and Morgan, G. (2010). Comparing child and adult development of a visual phonological system. *Lang. Interact. Acquisit.* 1, 67–81. doi: 10.1075/lia.1.1.05ort

Overton, W., and Jackson, J. P. (1973). The representation of imagined objects in action sequences: a developmental study. *Child Dev.* 44, 309–314. doi: 10.2307/1128052

Padden, C., Meir, I., Aronoff, M., and Sandler, W. (2010). "The grammar of space in two new sign languages," in *Sign Languages: A Cambridge Survey*, ed D. Brentari (New York, NY: Cambridge University Press), 570–592.

Padden, C., Meir, I., Hwang, S., Lepic, R., Seegers, S., and Sampson, T. (2013). Patterned iconicity in sign language lexicons. *Gesture* 13, 287–308. doi: 10.1075/gest.13.3.03pad

Perniss, P., Thompson, R. L., and Vigliocco, G. (2010). Iconicity as a general property of language: evidence from spoken and signed languages. *Front. Psychol.* 1:227. doi: 10.3389/fpsyg.2010.00227

Pfau, R., and Steinbach, M. (2006). "Modality-independent and modality-specific aspects of grammaticalization in sign languages," in *Linguistics in Potsdam 24*, eds H. Drenhaus, R. van de Vijver, and R. Vogel (Potsdam: Universitäts-Verlag), 5–98.

Piaget, J. (1952). *The Origins of Intelligence in Children.* New York, NY: International University Press.

Ravid, D., and Schiff, R. (2009). Morphophonological categories of noun plurals in Hebrew. *Linguistics* 47, 45–63. doi: 10.1515/LING.2009.002

Richie, R., Yang, C., and Coppola, M. (2014). Modeling the emergence of lexicons in homesign systems. *Topics Cogn. Sci.* 6, 183–195. doi: 10.1111/tops.12076

Sandler, W., Aronoff, M., Meir, I., and Padden, C. (2011). The gradual emergence of phonological form in a new language. *Nat. Lang. Linguist. Theory* 29, 503–543. doi: 10.1007/s11049-011-9128-2

Sandler, W., Meir, I., Padden, C., and Aronoff, M. (2005). The emergence of grammar: systematic structure in a new language. *Proc. Natl. Acad. Sci. U.S.A.* 102, 2661–2665. doi: 10.1073/pnas.0405448102

Schembri, A., Jones, C., and Burnham, D. (2005). Comparing action gestures and classifier verbs of motion: evidence from Australian Sign Language, Taiwan Sign Language, and nonsigners' gestures without speech. *J. Deaf Stud. Deaf Educ.* 10, 272–290. doi: 10.1093/deafed/eni029

Schick, B. (1987). *The Acquisition of Classifier Predicates in American Sign Language.* Doctoral Dissertation, Purdue University, West Lafayette, IN.

Slobin, D., Hoiting, N., Kuntze, M., Lindert, R., Weinberg, A., Pyers, J., et al. (2003). "A cognitive/functional perspective on the acquisition of 'classifiers,'" in *Perspectives on Classifier Constructions in Sign Languages,* ed K. Emmorey (Mahwah, NJ: Lawrence Erlbaum Associates), 271–296.

Spaepen, E., Coppola, M., Flaherty, M., Spelke, E., and Goldin-Meadow, S. (2013). Generating a lexicon without a language model: do words for number count? *J. Mem. Lang.* 69, 496–505. doi: 10.1016/j.jml.2013.05.004

Spaepen, E., Coppola, M., Spelke, E., Carey, S., and Goldin-Meadow, S. (2011). Number without a language model. *Proc. Natl. Acad. Sci. U.S.A.* 108, 3163–3168. doi: 10.1073/pnas.1015975108

Supalla, T. (1982). *Structure and Acquisition of Verbs of Motion and Location in American Sign Language.* Doctoral Dissertation, University of California, San Diego, CA.

Tang, G., Sze, F. Y. B., and Lam, S. (2007). "Acquisition of simultaneous constructions by Deaf children of Hong Kong Sign Language," in *Simultaneity in Signed Languages. Form and Function,* eds M. Vermeerbergen, L. Leeson, and O. A. Crasborn (Amsterdam: Benjamins), 283–316.

Thompson, R., Vinson, D. P., and Vigliocco, G. (2009). The link between form and meaning in American Sign Language: lexical processing effects. *J. Exp. Psychol. Learn. Mem. Cogn.* 35, 550–557. doi: 10.1037/a0014547

Tomasello, M., Striano, T., and Rochat, P. (1999). Do young children use objects as symbols? *Br. J. Dev. Psychol.* 17, 563–584.

van der Hulst, H. G. (1995). The composition of handshapes. *Trondheim Work. Papers* 23, 1–17.

Van Valin, R. (1993). *Advances in Role and Reference Grammar.* Amsterdam: Benjamins.

Vigliocco, G., Vinson, D. P., Woolfe, T., Dye, M. W., and Woll, B. (2005). Words, signs and imagery: when the language makes the difference. *Proc. R. Soc. Lond. Ser. B Biol. Sci.* 272, 1859–1863. doi: 10.1098/rspb.2005.3169

Werner, H., and Kaplan, B. (1963). *Symbol Formation: An Organismic-Developmental Approach to Language and the Expression of Thought.* New York, NY: Wiley.

Zwitserlood, I. (2003). *Classifying Hand Configurations in Nederlandse Gebarentaal [Sign Language of the Netherlands].* Utrecht: LOT (Netherlands Graduate School of Linguistics).

Zwitserlood, I. (2012). "Classifiers," in *Sign Language: an International Handbook,* eds R. Pfau, M. Steinbach, and B. Woll (Berlin: Mouton de Gruyter), 158–186.

Referential shift in Nicaraguan Sign Language: a transition from lexical to spatial devices

Annemarie Kocab [1,2]*, Jennie Pyers [2] and Ann Senghas [3]

[1] Department of Psychology, Harvard University, Cambridge, MA, USA
[2] Department of Psychology, Wellesley College, Wellesley, MA, USA
[3] Department of Psychology, Barnard College, Columbia University, New York, NY, USA

Even the simplest narratives combine multiple strands of information, integrating different characters and their actions by expressing multiple perspectives of events. We examined the emergence of *referential shift* devices, which indicate changes among these perspectives, in Nicaraguan Sign Language (NSL). Sign languages, like spoken languages, mark referential shift grammatically with a shift in deictic perspective. In addition, sign languages can mark the shift with a point or a movement of the body to a specified spatial location in the three-dimensional space in front of the signer, capitalizing on the spatial affordances of the manual modality. We asked whether the use of space to mark referential shift emerges early in a new sign language by comparing the first two age cohorts of deaf signers of NSL. Eight first-cohort signers and 10 second-cohort signers watched video vignettes and described them in NSL. Narratives were coded for lexical (use of words) and spatial (use of signing space) devices. Although the cohorts did not differ significantly in the number of perspectives represented, second-cohort signers used referential shift devices to explicitly mark a shift in perspective in more of their narratives. Furthermore, while there was no significant difference between cohorts in the use of non-spatial, lexical devices, there was a difference in spatial devices, with second-cohort signers using them in significantly more of their narratives. This suggests that spatial devices have only recently increased as systematic markers of referential shift. Spatial referential shift devices may have emerged more slowly because they depend on the establishment of fundamental spatial conventions in the language. While the modality of sign languages can ultimately engender the syntactic use of three-dimensional space, we propose that a language must first develop systematic spatial distinctions before harnessing space for grammatical functions.

Keywords: referential shift, narratives, spatial language, sign language, language creation

INTRODUCTION

Sign languages often exhibit a high degree of iconicity, compared to spoken languages, as signs and their referents exist in the same physical space (Taub, 2001). One arguably iconic component of sign languages is the use of distinct locations in signing space for grammatical purposes, such as locative marking and verb agreement (Klima and Bellugi, 1979; Supalla, 1982; Padden, 1983; Meier, 1987, 1990; Emmorey, 1996; Lillo-Martin and Meier, 2011). The present study explores the emergence of one class of grammatical devices, *referential shift* devices, that has been documented to include both lexical and spatial means to mark perspective changes (see Emmorey, 2002 for a review). We ask whether the earliest devices that emerge in Nicaraguan Sign Language (NSL) readily co-opted the iconic nature of space in the manual modality to mark shifts in reference.

When telling a story with multiple characters, a narrator must weave together a tapestry of information, integrating the perspective of the narrator with the perspectives of different characters to create a cohesive narrative. Signed and spoken languages alike employ a variety of referential shift devices to indicate multiple perspectives, and to mark when changes in perspective occur. To follow a narrative as it unfolds and to construct a mental representation of the event described, listeners rely on the narrator to provide information about the referents, their locations, their speech, and their actions. In spoken languages, narrators often use quoted speech to express different characters' points of view (Labov, 1972; Ochs, 1979; Schiffrin, 1981; Chafe, 1982; Tannen, 1982). English marks quoted speech with shifts in pronoun and tense, indicating a switched reference point for deixis. For example, in the sentence, "She said, 'I need more paint,'" the switch from the perspective of the narrator to the perspective of the character is marked syntactically by a shift from the third-person pronoun (*she*) in the matrix clause to the first-person pronoun (*I*) in the reported clause, and a shift from past (*said*) to present tense (*need*). Speakers can also rely on a shift in prosody to indicate quoted speech, changing intonation and voice quality to indicate something spoken by someone else (Clark and Gerrig, 1990). Because quoted speech is often not a faithful replication

of exactly what was uttered at the moment of the speech act, but rather a reconstruction of what was said, this part of a narrative is sometimes called *constructed dialogue* (Tannen, 1986). Constructed dialogue can express not only a character's speech, but also his or her thoughts and feelings. If we change the matrix verb *say* in the above example to either *be all* or *be like*, ("She was like, 'I need more paint'") the quoted clause now indicates the character's internal thoughts (Blyth et al., 1990).

Like spoken languages, sign languages employ constructed dialogue, with a shift in deictic perspective, to express a character's speech and thoughts. Additionally, the manual modality allows signers to express a character's actions using a device known as *constructed action* (Liddell and Metzger, 1998). When representing actions, signers use their own bodies (*embodiment*) including the face, torso, and arms, to convey information about different characters and their actions. For instance, consider a story about a woman who enters a room and accidently lets the door close in the face of someone trying to enter behind her. A signer narrating this story could embody the characters' actions and reactions, first enacting closing a door with an expression of naïve ignorance, and then enacting bumping into a door and adopting an expression of surprise.

Importantly, the effective representation of multiple perspectives, through both constructed dialogue and constructed action, depends on clear and unambiguous marking of when there is a shift in reference. The narrator must clearly introduce the different characters, and when describing their speech, thoughts, and actions, unambiguously indicate which character's speech, thoughts, and actions are being expressed. If a narrator does not mark perspective shifts clearly, the listener may mistakenly attribute all of the speech, thoughts, and actions to a single character. As such, coherence and clarity in a narrative rely on the narrator's systematic use of lexical and grammatical cues to indicate who the referents are and when a shift in perspective occurs.

Referential shift devices similar to those found in spoken languages, such as pronominal shifts and pauses, have been identified in several sign languages, including American Sign Language (ASL), British Sign Language, Danish Sign Language, Swedish Sign Language, and South African Sign Language (Loew, 1984; Shepard-Kegl, 1985; Padden, 1986, 1990; Liddell, 1990; Lillo-Martin and Klima, 1990; Meier, 1990; Engberg-Pedersen, 1993; Poulin and Miller, 1995; Aarons and Morgan, 2003; Janzen, 2004;

Cormier et al., 2013). Sign languages also leverage non-manual elements, signaling referential shifts through breaks in eye-gaze, head tilts, facial expression, and a *body shift* (Padden, 1986)[1] from a neutral position to a specified spatial location associated with the referent (Engberg-Pedersen, 1993; Cormier et al., 2013). These devices differ from embodiment, where signers use the whole body or a part of the body to represent a particular character's body or its parts, in that these non-manual movements of the body signal a perspective change rather than convey information about a character's actions.

Referential shift devices in sign languages can be broadly categorized into two types of devices: *lexical* and *spatial*. The first category employs spatially neutral lexical devices to indicate a shift in referent. For instance, signers can assign a referent a *lexical label*, such as WOMAN[2]. Using that label, narrators can then introduce and re-introduce the character associated with the label, allowing the listener to understand when a shift to that character's perspective has occurred or is about to occur (for a review, see Cormier et al., 2013). An example of the use of the lexical label in NSL to indicate a shift in referent can be seen in **Figure 1**, where the signer represents the constructed actions of two different characters, one who lifts books down to the other, who receives them. Immediately before the second instance of constructed action, the signer produces the sign WOMAN to indicate a shift in reference to the new character who receives the books. This referential shift device indicates that the agent of the receiving action is a different character from the agent of the previous lifting action.

---

[1]This type of role-shifting, where there is a change in body position to indicate a shift in perspective, has been termed *contrastive role-shifting*. Padden (1986) argues that in contrastive role-shifting, at most only two roles can be contrasted. In the case of narratives with more than two referents, the third referent is introduced in a different (subordinate) role-shifting structure. The subordinate structure (the third referent and one of the original two referents) is contrasted with the initial contrastive role-shift structure (first and second referent). Importantly, the contrastive role-shifting structures are still associated with separate spatial locations.

[2]English glosses for signs appear in SMALL CAPS.

[3]Following the conventions in sign linguistics literature (see Emmorey, 2002; Cormier et al., 2013), a token of constructed action is indicated with "CA:" followed by a description in lower-case letters. The concepts described with CA are italicized. When a referent is specified, CA:x is used, where x is the referent, with angled brackets placed at the beginning and end of the CA.

| BOOK | _{CA:woman} *CA:lifting-book* | WOMAN | _{CA:woman} *CA:receiving-book* |
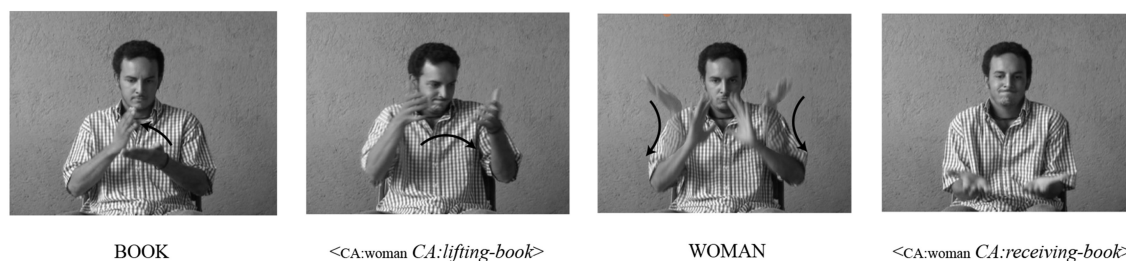
FIGURE 1 | An example of a spatially neutral *lexical label* as a referential shift device to mark perspective change in NSL. The sign WOMAN in the third panel, produced in neutral space, is a lexical label marking a shift to the perspective of a new character[3].

Prior work on referential shift in NSL has noted a second lexical device that has not yet been documented in other sign languages: a point to the chest (Pyers and Senghas, 2007). Using this device, signers point to themselves, explicitly indicating that they are about to take on the role of a character. This *point-to-chest* is sometimes, but not always, followed by a lexical label or a description of the character whose perspective the signer is about to adopt (e.g., the person with the books). The point-to-chest used in NSL in **Figure 2** is distinct from the first-person pronoun used with referential shift in ASL in that the point-to-chest is produced before the shifted construction, signaling that the signer is about to change to the perspective of a particular character and construct the actions of that character, with a neutral positioning of the torso and shoulders, while the first-person pronoun in ASL is produced after a referential shift has been established, indicating self-reference by a character.

In contrast to lexical devices, *spatial devices* capitalize on the visual-spatial nature of sign languages and the ability of the signer to associate referents with locations in the three-dimensional signing space in front of the signer. For example, in many mature sign languages, specific locations in signing space are first associated with nominal signs. Once referents are associated with unique locations, the signer can anaphorically refer back to these locations, using direction of eye gaze, pointing, or a body or head shift. Examples of spatial devices in NSL are shown in **Figures 3–5**. In **Figure 3**, the signer uses a body shift along with his constructed action sequences to indicate the switch in perspective between two characters, one who draws on a whiteboard and another who then erases it.

Another spatial device to mark changes in perspective is an *indexical point to space* (**Figure 4**). Here the signer locates referents in different locations in the signing space, then points to one
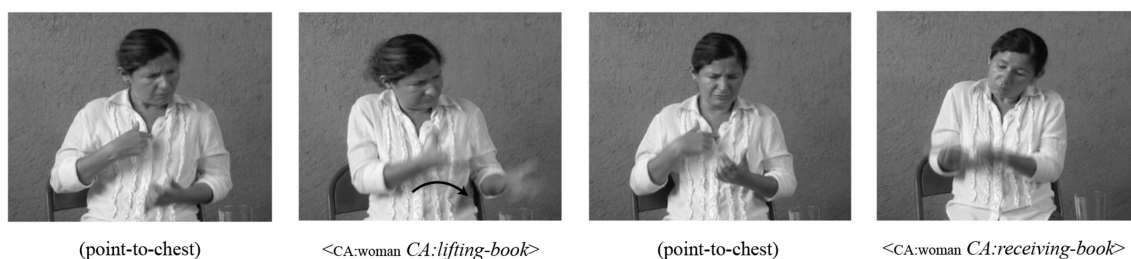
FIGURE 2 | An example of the *point-to-chest* to mark perspective change in NSL. In the first panel, the point-to-chest, produced with the torso and shoulders in neutral position, indicates a shift to the perspective of the first character. In the third panel, a second point-to-chest marks a shift to the second character.

FIGURE 3 | An example of *body shift* as a referential shift device to mark perspective change in NSL. The movements of the torso to the signer's left in the second panel, and to his right in the fourth panel, indicate referential shifts from one character to another.
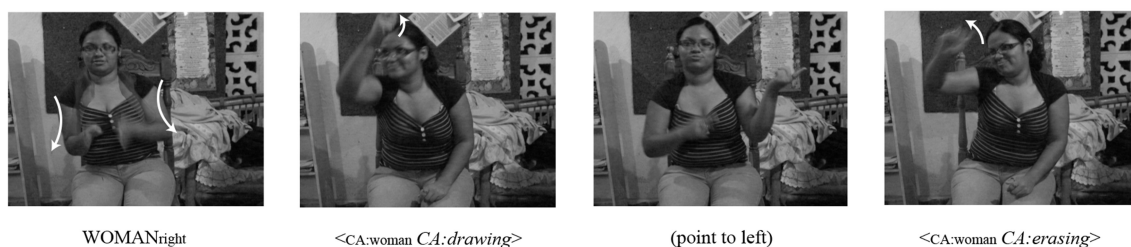
FIGURE 4 | An example of an *indexical point-to-space* as a referential shift device to mark perspective change in NSL. The signer points to her left in the third panel to indicate a referential shift from one character to another.

of these locations before engaging in constructed action in order to indicate a change in reference to the referent associated with that location in space.

A third spatial device in NSL is the *spatially modulated lexical label* (**Figure 5**). Here a lexical sign is produced in a specific location in the signing space, rather than in the neutral area in front of the signer's body. This device, like other spatial devices, can be used at first mention of a referent to establish a character in a narrative, or in later mentions in a narrative to shift reference to that character.

The use of space to indicate the locations of and relations among referents derives from the iconic relationship between the spatial representations in signing space and the true locations of the referents in the world (Emmorey and Reilly, 1995; Engberg-Pedersen, 1995; Taub, 2001). This iconic use of space is prevalent across mature sign languages, and has also been observed in gestural communication systems, called *homesigns*, that are developed by deaf children with their hearing family members when sign language is not available (e.g., Goldin-Meadow and Mylander, 1990; Engberg-Pedersen, 1993, 1995; Emmorey and Reilly, 1995; Coppola, 2002; Morgan et al., 2008). Bosworth and Emmorey (2010) suggest that the prevalence of iconicity may stem from the gestural origins of sign languages, perhaps due to the functional pressure for clarity and ease of communication. As such, during the emergence of a new sign language, when gestures and homesigns are reorganized into a structured language, one might expect to see creators of a new sign language

readily avail themselves of the iconic nature of space to structure narratives.

Because the signer's body can represent multiple characters and their different perspectives, as well as the signer's own perspective as the narrator, signers of mature sign languages generate different types of formats in their iconic representations of real-world spatial relations; these format types differ in perspective. One spatial format, *diagrammatic* space, situates the signer outside of the event, describing it from an observer's point of view (as in **Figure 5**). A second spatial format, *viewer* space, locates the signer within the event itself, describing it from an experiencer's point of view (as in **Figure 3**), and expressing any spatial relations relative to the character's first-person perspective (Emmorey and Falgier, 1999) (see **Figure 6**)[4]. Using these spatial representations, signers convey information about objects and characters, their locations, and spatial relations.

Further, the signer has two options for representing the signing space (see **Figures 7**, **9**): the signer can use the front-back axis or the left-right axis (Padden, 1986). When using left-right spatial contrasts, signers set up a referent on one side of the signing space (either left or right) and contrast it with a second referent set up on the opposite side of the signing space (Emmorey, 2002), as in **Figure 5**. With front-back spatial relations, signers tend to

---

[4]Other terminology for these two spatial formats include observer perspective vs. character perspective (Perniss, 2007), depictive vs. surrogate space (Liddell, 2003), and narrator vs. protagonist perspective (Slobin et al., 2003).
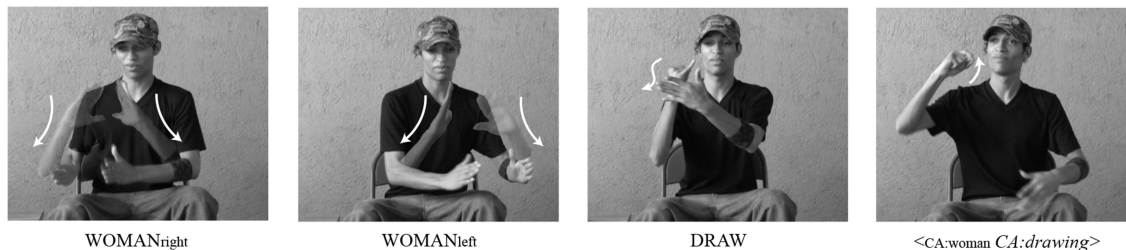
FIGURE 5 | An example of *spatially modulated lexical labels* in NSL. In the first panel, the signer produces the sign WOMAN to his right, thereby associating the first woman with that location. In the second panel, he produces the sign WOMAN to his left, associating the second woman with that second location.
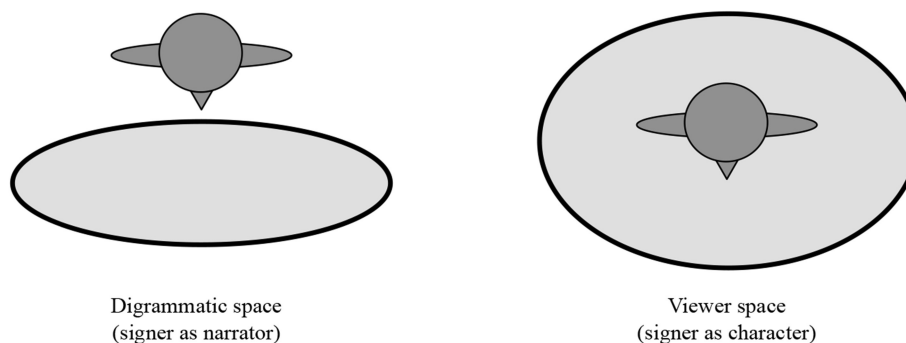
FIGURE 6 | Schematics of *diagrammatic space* and *viewer space* formats. The dark gray figures represent the signer viewed from above, and the light gray oval represents the area of the signing space in which spatial signs are produced.

FIGURE 7 | Schematics of (A) front-back and (B) left-right spatial layouts within a diagrammatic space format.

WOMAN   (point to front)   WOMAN   CA:woman <CA: drawing>

FIGURE 8 | An example of a front-back spatial layout within diagrammatic space in NSL. The signer's point to a spatial location in front of the signer in the second panel associates that spatial location with the character of the second woman.

FIGURE 9 | Schematics of (A) front-back and (B) left-right spatial layouts within a viewer space format.

embody an animate character, and use the body as a locus with respect to which other characters or objects are assigned locations, in front of or behind the signer, as in **Figure 8** (Emmorey, 2002; Perniss, 2007).

In mature sign languages, the preferred layout can depend on format. In ASL, diagrammatic space is typically used with left-right spatial contrasts, where the signer is narrating from the perspective of an observer, and contrasting referents are set up in front of the signer in the signing space. When using a viewer space format, where the signer is narrating from the perspective of a character, both left-right and front-back spatial contrasts can be used (Emmorey, 2002). These patterns have been found in ASL, a well-documented and mature sign language; we might expect to see different patterns of spatial layouts and formats cross-linguistically, and in a younger, emerging sign language.

Though the placement of referents in signing space is often iconic and derives its structure from the relative locations of objects in the world, spatial signing is not necessarily transparent (Emmorey and Reilly, 1995). As a signer narrates a story, the listener must construct a mental representation of the event, relying on the spatial information presented by the signer. As the narrative unfolds, the listener must continually map new information onto the developing mental image (Givón, 1995; Gernsbacher, 1997). Thus, narrative comprehension in sign languages is partly dependent on the signer's establishment and maintenance of the distinct spatial relationships among the referents throughout the

narrative, particularly for short narratives of a specific event or situation[5], such that the listener knows which referent's speech and actions are being represented (Winston, 1991). With such spatial consistency a subtle shift of the body or head relative to a spatial location associated with a referent can be sufficient to communicate a change in perspective (Emmorey, 2002).

To appreciate why setting up referents in spatial locations is crucial for narrative coherence, consider a signer who uses the body to represent multiple perspectives, via constructed action, but does not overtly mark when perspective changes occur. In such a situation, the listener may correctly understand that the narrator is telling a story about, say, one man who was walking and another man who was eating. Alternatively, the listener could easily misinterpret the account to be that a single man was walking and eating. Lexical devices could disambiguate which characters performed specific actions, without providing spatial information that reveals how the characters are located relative to one another. However, for spatial devices to be understood correctly, the signer must consistently map the different referents to contrastive locations in the signing space, across multiple signs, including constructed action sequences. As such, signers sometimes use non-spatial explicit referential shift devices such as lexical labels alongside spatial devices such as a body shift to make the referent more salient than it would be with the spatial device alone (Cormier et al., 2013).

The present study examined the origins of these complex grammatical systems. We follow the development of spatial and lexical grammatical devices for marking referential shift in an emerging sign language. By examining the early stages of a young sign language, we asked whether the richness of the spatial iconicity prevalent in the manual modality motivates early emergence of spatial devices in a new sign language.

The language under consideration emerged over the past four decades in Managua, the urban capital of Nicaragua. Prior to the 1970s, there was no established sign language in use in Nicaragua, but special education reforms in the late 1970s and early 1980s brought about drastic changes. With the opening of a primary school for special education, followed by a vocational center, deaf children and adolescents were able to socialize in greater numbers than ever before, giving rise to the birth of a new language (Kegl and Iwata, 1989; Polich, 2005). An initial group of fifty signers passed on their developing language to waves of new children entering the community each year, who, in turn, continued to add to the language's complexity and development (Senghas, 1995; Senghas and Coppola, 2001). By comparing the language of that initial *first cohort* of signers to that of those who entered in the language's second decade, the *second cohort*, we can see how the language has changed and grown.

The recent emergence of NSL offers the opportunity to examine when referential shift devices emerge in a language, which devices emerge first, and whether early-emerging devices differ from later-emerging devices in their use of space. We investigated the distribution and frequency of use of specific referential shift devices in NSL, and the pattern of device use between cohorts. If

the creators of a new sign language can immediately harness the iconic nature of three-dimensional space, one might expect to see early emergence of spatial devices to consistently mark referential shift in the first cohort of Nicaraguan signers.

MATERIAL AND METHODS

PARTICIPANTS

Eighteen deaf Nicaraguan signers participated in the study, ranging in age from 21.4 to 40.0 years. We grouped participants into two cohorts according to the year they were first exposed to NSL when they entered the primary school for special education (Senghas, 1995): 8 first-cohort signers (5 M, 3 F, $M_{age} = 33.1$ years) were exposed to NSL before 1986, and 10 second cohort signers (6 M, 4 F, $M_{age} = 24.0$ years) were exposed to NSL between 1986 and 1990. All participants were exposed to NSL by the age of 6 (Cohort 1 mean age of exposure: 4.6 years; Cohort 2 mean age of exposure: 4.0 years). All participants gave consent to participate and be videotaped as part of this study, and all were paid for their participation. The research protocol was approved by the Barnard College Institutional Review Board for the Protection of Human Subjects in Research and by the Wellesley College Psychology Department Research Ethics Committee.

MATERIALS

Participants were shown six video vignettes, presented as QuickTime movies on a laptop computer. The vignettes were 10–30 s in duration, depicting simple events that included two or three characters performing straightforward actions with no dialogue (**Table 1**). The vignettes were designed to eliminate the need for participants to make inferences about internal states to understand the actions depicted. Accordingly, the actions performed by the characters did not imply any hidden beliefs or intentions. For example, signers could (and often did) produce descriptions of Video #3 like, "The woman on the left was drawing on a whiteboard" (see **Table 1**). Of course, participants' responses could include descriptions of the characters' emotions or mental states; narrative information that goes beyond what is said is often a part of constructed dialogue.

PROCEDURE AND CODING

Participants watched each vignette and were instructed in NSL to "describe [to another signer from their cohort] what you saw." Participants were permitted to watch the movies as many times as they liked. Narratives were videotaped (30 fps) for coding offline. Elicited narratives were coded by the first author, who is a fluent signer of ASL with 6 years of research experience with NSL, for (1) average length of the narrative, (2) proportion of perspectives represented, (3) use of space to assign spatial locations to referents, and the spatial format and spatial layouts used, (4) overt marking of referential shifts, and (5) types of referential shift devices used.

The average length of each narrative was computed in seconds. The counter started from the moment the signers lifted their hands until they dropped their hands, signaling the end of the narrative. The length of the narrative was calculated to check whether a greater number of perspectives represented might be a simple consequence of longer narratives.

---

[5]In longer narratives, the signer may reset the spatial locations associated with referents, using different loci at different points in the story (van Hoek, 1992).

Because the number of characters in the vignettes varied (see **Table 1**), we calculated the proportion, rather than the sum, of perspectives represented. This proportion was defined as the sum of perspectives represented by the signer divided by the total number of possible perspectives included in the narrative (the number of characters in the narrative plus the signer's perspective as the narrator; **Table 1**). Signers were coded as having represented the narrator's perspective if the signer included commentary or descriptive information from the signer's own perspective. Signers were coded as having represented a given character's perspective if the signer engaged in constructed action representing that character's perspective, such as imitating the facial expression, body posture/orientation or actions of the character, with a maximum score of 1 for each character, regardless of how many times that character's perspective was represented. For instance, if the signer was describing Video #3 and imitated the action of erasing a whiteboard, that sequence was coded as constructed action, representing the perspective of the character who did the erasing. Crucially, credit was given for representing a character's perspective even if the signer did not grammatically mark the perspective shift. Indeed, to someone naive to the video stimuli, many of the representations of the actions of multiple characters were produced without such marking, and could consequently be misinterpreted as multiple actions by a single character. For example,

a listener might interpret a signed narrative with instances of constructed action of drawing and erasing either as a single character who is drawing and then erasing her own picture, or as two different characters, one who draws a picture and one who erases it. We included this measure of the number of perspectives represented to capture only whether participants explicitly encoded and expressed the actions of the different characters.

In addition, we coded whether signers assigned spatial locations to the referents at the beginning of each narrative, and whether they implemented a diagrammatic or viewer spatial format. We further looked at whether signers used front-back or left-right spatial layouts when assigning spatial locations to referents.

Signed narratives were also coded for the use of five types of referential shift devices that have been previously observed in NSL (Pyers and Senghas, 2007): *lexical label*, *point-to-chest*, *body shift*, *indexical point-to-space*, and *spatially modulated lexical label* (see **Table 2**). *Lexical label* and *point-to-chest* are lexical, non-spatial devices that indicate shifting to a particular character's perspective (see **Figures 1**, **2**). *Body shift*, *indexical point-to-space*, and *spatially modulated lexical label* are spatial devices that associate physical locations in the signing space with particular referents (see **Figures 3–5**). The use of a referential shift device was coded as positive whenever the signer employed the device before

Table 1 | Stimulus characteristics.

Video #	Description	Number of perspectives, including the narrator's
1	Two women; one seated on a stool; the other approaches her and dumps a wastebasket full of paper over her head	3
2	Two women; one gives the other three stacks of books to hold	3
3	Two women; one draws on a whiteboard while the other erases her drawing	3
4	Two women; one throws a ball to the other, who holds a stack of papers, and drops the stack to catch the ball	3
5	Two women; both race to sit on a stool; one succeeds and knocks the other to the floor	3
6	Three women; one is standing inside a room, a second knocks on the door and the first lets her in; a third woman attempts to enter the room as the door is closing, but the door closes on her face, seen through a glass window in the door	4

Table 2 | Referential shift devices.

Type of device	Referential shift device	Description
Lexical	Lexical label	The signer's use of a spatially neutral lexical label, such as a noun, to indicate a change in reference. The most common lexical label used was WOMAN, as women were featured in all of the character roles.
	Point-to-chest	The signer's use of a point to the chest, with the torso and shoulders in a neutral position, to indicate a change in reference.
Spatial	Body shift	The signer's use of a shoulder/body shift in physical space to the right or left from a central axis to indicate a change in reference.
	Indexical point-to-space	The signer's use of a point to a location in space to indicate a change in reference to the referent associated with that location.
	Spatially modulated lexical label	The signer's use of a lexical label produced in a particular location in space to indicate a change in reference to the referent associated with that location.

introducing or re-introducing a character's or the narrator's perspective, and when shifting between perspectives. We coded the use of these devices only if they were used to indicate a shift in the referent before the signer engaged in constructed action, using his or her body to represent the perspective of the character, expressing the character's actions or feelings.

We analyzed only whether a particular device was present in each narrative, not the frequency with which that particular device appeared in the narrative. For instance, if a signer used all five referential shift devices while describing one vignette, that narrative would receive a 1 for each type of referential shift device. In cases where the signer used multiple devices to mark a single referential shift, each type of device used was coded as present.

RESULTS

First, we considered the average length of signers' narratives in the two age cohorts to determine whether any difference in the number of perspectives represented might be a reflection of the amount of time spent describing the events. The average signing time did not differ significantly between cohorts [Cohort 2: 15.50 s, $SD = 5.55$; Cohort 1: 12.50 s, $SD = 4.39$, $t_{(16)} = 1.29$, $p = 0.22$, two-tailed]. Each vignette contained multiple characters, and we measured the proportion of characters whose perspectives were represented or mentioned by signers in their narratives. Both cohorts expressed the majority of available perspectives in each narrative (see **Table 3** for means and standard deviations). There was a marginally significant difference between the two cohorts in the number of perspectives represented [$t_{(7.30)} = 2.33$, $p = 0.07$, two-tailed, adjusted for unequal variances], although a Levene's test showed that the first cohort signers were significantly more variable in their performance ($F = 13.75$, $p < 0.01$). This difference between cohorts in the number of perspectives represented was driven primarily by the inclusion of explicit marking of and shifts between the perspective of the narrator and of a character. Since the narratives were not based on first-hand accounts, and were always told from the perspective of an outside observer (that is, the signer as the narrator), we conducted an additional analysis comparing the proportion of perspectives represented aside from that of the narrator, and

found no significant difference between cohorts [**Table 3**, $t_{(8.96)} = 1.52$, $p = 0.20$, two-tailed, adjusted for unequal variances, $F = 9.71$, $p < 0.01$]. According to this analysis, signers from both cohorts were similarly able to represent the perspectives of the characters in their narratives.

Next, we examined whether signers differed in their assignment of spatial locations to referents at the beginning of each narrative. Due to the categorical nature of the dependent variables, and because we were conducting between-subjects analyses, we used logistic mixed effects regression with item (video) and subject as random effects, where the use of space to assign spatial locations to referents and the absence of use of space were entered as 1 and 0, respectively. In the model, we looked at the effect of cohort on whether signers assigned spatial locations when describing the videos. The predictor variable, cohort, was coded such that the first cohort (signers who entered the community prior to 1986) represented the baseline (the intercept). Positive and negative coefficients are interpreted with respect to this intercept value, where a positive coefficient (β) represents an increase in the likelihood of the second cohort using the dependent variable of interest, and a negative coefficient represents a decrease. Accordingly, we report the coefficient representing the second cohort, indicating the difference from the first cohort, followed by Wald's z-score.

There was no difference between the two cohorts in how many narratives included the use of space for assignment of spatial locations to referents (β = 3.52, $Z = 0.00$, $p = 1.00$). We then considered whether the two cohorts differed in their use of diagrammatic and viewer space in their narratives. Second-cohort signers used diagrammatic space significantly more than first-cohort signers (β = 3.10, $Z = 3.84$, $p < 0.001$), but signers from the two cohorts did not differ in their use of viewer space (β = 3.41, $Z = 0.00$, $p = 0.99$). Note that signers can and did use both spatial formats within a single narrative. We then considered the frequency with which the two types of spatial layouts, front-back and left-right, were used by signers from each cohort in their narratives. The two cohorts did not differ in their use of front-back spatial relations overall (β = −0.47, $Z = −0.67$, $p = 0.50$), but did differ significantly in their use of left-right distinctions (β = 3.60, $Z = 4.69$, $p < 0.001$). In other words, signers from the two cohorts did not differ in how often they used space to assign locations to referents overall, nor did they differ in their use of front-back spatial relations when taking on the perspective of a character in the event (locating the other character in the space in front of them), but they did differ in their use of diagrammatic space (locating the narrator outside the event as an observer) and their use of left-right spatial relations for their referents.

Finally, we analyzed the frequency and type of referential shift devices used in the narratives. The data were submitted to a logistic mixed effects regression with item (video) and subject as random effects, and use of a specific referential shift device was assigned a 1 and the absence of that referential shift device was assigned a 0. Second-cohort signers used referential shift devices to explicitly mark a shift in perspective in significantly more narratives than did first-cohort signers (see **Table 4** for means and standard deviations, β = 3.17, $Z = 2.75$, $p < 0.01$). Crucially, we observed a difference between the two cohorts in how consistent

Table 3 | Proportion of perspectives represented, and proportion of narratives in which signers set up referents in space, and used front-back and left-right spatial layouts, by each cohort.

	Cohort 2 ($N = 10$)	Cohort 1 ($N = 8$)
	Proportions (*SD*)	Proportions (*SD*)
Perspectives (narrator's and characters')	0.98 (0.04)	0.82 (0.22)
Perspectives (characters' only)	0.99 (0.02)	0.96 (0.06)
Assigned spatial locations	0.98 (0.05)	0.67 (0.43)
Diagrammatic space	0.47 (0.22)	0.21 (0.19)
Viewer space	1.00 (0.00)	0.96 (0.12)
Front-back	0.50 (0.18)	0.63 (0.21)
Left-right	0.88 (0.14)	0.52 (0.26)

Standard deviations given in parentheses.

Table 4 | Proportion of narratives that included each referential shift device, by each cohort.

Referential shift device	Cohort 2 (*N* = 10)	Cohort 1 (*N* = 8)
	Proportions (*SD*)	Proportions (*SD*)
Perspectives grammatically marked	0.95 (0.11)	0.67 (0.30)
Lexical Devices	0.58 (0.50)	0.50 (0.51)
Lexical label	0.52 (0.50)	0.38 (0.49)
Point-to-chest	0.20 (0.31)	0.19 (0.29)
Spatial Devices	0.87 (0.34)	0.27 (0.45)
Indexical point-to-space	0.38 (0.32)	0.06 (0.09)
Body shift	0.75 (0.29)	0.17 (0.24)
Spatially modulated lexical label	0.38 (0.49)	0.10 (0.31)

Standard deviations given in parentheses.

they were as a group in marking referential shift. Five first-cohort signers used referential shift devices in at least five of the six narratives, while the remaining three first-cohort signers used them in three or fewer narratives, that is, half the time or less. In contrast, nine of the 10 second cohort signers in our study used referential shift devices in at least five narratives, and the remaining second-cohort participant used them in four of the six narratives.

In investigating the types of devices used to mark referential shift, we found that second-cohort signers used significantly more spatial devices ($\beta = 3.38$, $Z = 5.57$, $p < 0.001$) but not more lexical devices ($\beta = 0.36$, $Z = 0.76$, $p = 0.45$) than first-cohort signers. There was no significant difference between cohorts in the use of neutral lexical labels as a device ($\beta = 0.58$, $Z = 1.46$, $p = 0.14$) nor point-to-chest ($\beta = -0.75$, $Z = -0.36$, $p = 0.72$). We next looked at whether there were cohort differences in the use of the different types of spatial devices. Compared to the first-cohort signers, second-cohort signers used significantly more spatially modulated lexical signs ($\beta = 1.78$, $Z = 2.79$, $p < 0.01$), indexical points-to-space ($\beta = 2.71$, $Z = 2.77$, $p \leq 0.01$) and body shifts ($\beta = 3.96$, $Z = 4.05$, $p < 0.001$).

DISCUSSION

The recent emergence of a sign language in Nicaragua offers us the opportunity to capture the creation and development of new grammatical devices. We followed the emergence of referential shift devices over the first two sequential age cohorts of NSL, paying particular attention to the degree to which signers leveraged the iconic use of space for this function. There are reasons to expect referential shift to take advantage of spatial iconicity from the outset. Spatial devices for referential shift have been found in many mature sign languages, and may turn out to be a sign language universal. Furthermore, if a sign language already incorporates highly embodied, iconic representations within constructed action to depict the behaviors of characters in a narrative, it seems a natural first step to refer to the relative spatial locations of those characters to mark a shift in perspective from one character to another.

We observed that both the first and second cohorts of signers of NSL easily represented multiple characters' perspectives, readily switched back and forth among these perspectives, and did not differ in the number of character perspectives represented in their narratives. Previous work has documented delays in false-belief understanding in first-cohort signers (Pyers and Senghas, 2009), which had made us sensitive to the possibility that first-cohort signers might not effectively represent the different perspectives within a story. We found, however, that this was not the case; members of both cohorts, with equal frequency, encoded and represented the different characters' roles in their narratives.

Where the cohorts differed was in the use of devices to explicitly mark the shift from one perspective to another. This grammatical marking of referential shift was significantly greater in the second cohort. Both cohorts expressed perspective change using referential shift devices at least some of the time, suggesting that the seeds for linguistic marking of perspective emerged early in the language. However, the first cohort was both less frequent and more variable in their marking than the second. While three of the eight first-cohort signers marked referential shifts in half or fewer of their narratives, it was rare for a second-cohort signer to perform a shift without explicitly marking it. The consistency observed across the second-cohort participants suggests that over the late 1980s, while they were still young, NSL became increasingly stable in the marking of perspective changes in a narrative.

Where referential shift did appear in the signing of the first cohort, it was primarily as a non-spatial, lexical device; spatial marking, though present, was used far less. In contrast, second-cohort signers used spatial devices significantly more than the first-cohort signers did. This pattern of findings suggests that the use of space to mark referential shift was somewhat slow to emerge, relative to lexical devices.

The later emergence of spatial devices to grammatically mark referential shift does not appear to be due to the lack of a productive use of spatial layouts in general, throughout the language. Signers from both cohorts used the signing space in a concrete, iconic way to assign spatial locations to referents, along both the front-to-back and left-to-right axes, in about half of their narratives. This frequency suggests that both kinds of layouts, and the explicit use of the three-dimensional signing space, have been available since the earliest years of NSL. As we move from the first to the second cohort, the assignment of referents to locations to the left and right increased. Interestingly, we did not see a similar increase in the use of the front-to-back axis. Evidently, as the language matured, the balance between the two layouts changed, favoring differentiation along the left-to-right axis, at least for this function.

Along with the change in spatial layout, we observed changes in the nature of the spatial format applied in the narratives. Both cohorts readily described events using a character's perspective, even multiple characters' perspectives, in viewer space. That is, signers from both cohorts adopted the perspective of a character within their constructed action utterances. However, the use of diagrammatic space to frame the event from the perspective of an outsider, here the narrator, increased across cohorts, occurring in less than a quarter of the first cohort's and about half of

the second cohort's narratives. We suggest that these changes—the increase in the use of differentiation along the left-to-right axis, and the increase in the use of diagrammatic space to structure the narrative—follow from other changes in the language, specifically, (1) the establishment of conventions for conveying left-right spatial contrasts (*to the left of*, *to the right of*) and (2) the development of more complex story structures, framed at the level of a third-person narrator.

THE EMERGENCE OF LEFT-RIGHT SPATIAL CONTRASTS

Despite its apparent iconicity, the use of space to the left and right of a signer to convey the concept of physical spatial contrasts, such as *left of* and *right of,* is not automatic or transparent, and was not available at the outset of NSL. Descriptions of left-right relations are grounded in real-world space, and the mapping from real-world-space to signing-space can be ambiguous. Left-right contrasts present a particular challenge, because they fall along an axis of symmetry—the left side of the body is symmetrical to the right side—and because perspective differs from one persons' viewpoint to another. This combination makes them more subject to ambiguity than up-down and front-back contrasts. The convention in ASL and many other mature sign languages is that left-right spatial contrasts are typically described from the viewpoint of the signer (Emmorey, 1996; Pyers et al., 2008). Previous work on NSL has shown that second-cohort signers introduced consistency in the use of spatial language, systematically marking left-right spatial relationships, and linguistically distinguishing among contrastive locations within the signing space. The older, first-cohort signers do use the signing space to describe spatial relationships, but do not do so systematically, resulting in ambiguous spatial descriptions (Pyers et al., 2010). Specifically, first-cohort signers might use the same spatial locations to describe objects to their left as for objects to their right, while second-cohort signers would use distinct locations to the left and right side of the signing space to convey the relative locations of objects to the left and right side in the real world. Evidently this spatial contrast took some time to develop in NSL, and did not conventionalize until the first cohort had already reached adolescence. Once the language had established conventions for left-right spatial contrasts in descriptions of physical space, signers readily applied this distinction in devices marking abstract reference. In that sense, the spatial referential shift devices developed quickly.

THE DEVELOPMENT OF THE NARRATOR WITHIN THE NARRATIVE

Conventions at the level of discourse structure similarly took some time to emerge, and did not necessarily arise automatically once more local devices for sentence structure were available. In related work on the emergence of NSL, the use of devices for building discourse cohesion increased as the language was passed from the first to the second cohort. These developments included an increase in the range and frequency of devices for explicitly marking grammatical subjects (Coppola et al., 2013) and the development of anaphoric uses of pointing (Coppola and Senghas, 2010). Consequently, signers could manage narratives more explicitly at a meta-level, introducing and referring back to characters unambiguously. The introduction and maintenance of

characters within a story is often performed at the discourse level of the narrator, across sentences. Increasing grammatical specificity at the local level (e.g., distinguishing subjects and objects) likely enabled more complex narrative structure, which in turn created a context in which any referential devices needed to be applied consistently to effectively maintain reference in longer utterances, including across sentence boundaries. That is, the development of more complex narratives may have created pressure to express multiple distinct perspectives unambiguously, and to explicitly distinguish the narrator's perspective from the characters' perspectives in a story. Lexical and spatial devices being used to disambiguate reference might have then been taken up as grammatical markers of referential shift.

SIMILARITIES TO OTHER EMERGENT SYSTEMS

Work on the development of verb agreement in Israeli Sign Language (ISL), a language around 75 years old, shows a similar pattern of a shift from use of front-back space to left-right space as the language develops (Meir, 2012). The earliest inflected forms of agreeing verbs in ISL are produced on the front-back axis, where the argument is associated with a spatial locus in front of the signer, and the directionality of the sign moves from the signer's body to the locus. In later forms, arguments are associated with spatial locations to the left or right of the signer, and inflected signs are produced along a left-right axis, originating from the signer's body or a spatial location off the body and ending at the locus associated with the object or recipient argument.

In contrast, Al-Sayyid Bedouin Sign Language (ABSL), an emergent village sign language that is approximately the same age as ISL, has not yet developed ways of using space in this grammatical way, and lacks a spatial verb agreement system (Meir et al., 2007). This later, or absent, grammatical use of the signing space may be indicative of differences between deaf community sign languages, like ISN and NSL, and village sign languages like ABSL (see Meir et al., 2010). Village sign languages typically have a smaller number of deaf members—in this case, ABSL has about a tenth the number of deaf members as NSL—with a correspondingly higher degree of shared information (Sandler et al., 2005). The resulting greater intimacy within the community may put less functional pressure on the language to make the kinds of grammatical distinctions we document in NSL. We would predict that any future development of a grammatical use of space in ABSL would include viewer-space perspective and front-back contrasts emerging first, and diagrammatic-space perspective and left-right contrasts appearing later.

Note that the changes we explore here do not represent a simple increase or decrease in spatial iconicity, but rather, a change in the nature of the spatial mappings that different grammatical devices exploit. We can identify two types of iconicity in use, which map on to different aspects of events. The first type, which includes enactment, depicts the actions of agents from the agent's own perspective. The second type, which includes diagrammatic-space depictions and spatial body shifts, depicts the actions and locations of agents from a narrator's perspective. These two types of iconicity use mutually incompatible mappings. For example, through enactment, a signer can faithfully replicate the behaviors of the referent, using the movement of the signer's body to

represent movement of a referent's body. But when the signer moves the body to indicate a shift of reference, that body movement no longer maps to the movement of the referent's body. There is no corresponding movement, by any referent, that occurs in the actual event. The first type of iconicity, enactment, is highly effective in portraying the actions of a single character, but is limited in its capacity to depict other components of an event, such as other referents and their actions. Conversely, spatial iconicity captures the relationship among the components of the event. Once you have both types of iconic representations in use in a language, grammatical devices are necessary to effectively switch back and forth between the two.

PARALLELS TO ACQUISITION

In some ways, the developments we have documented in the emergence of NSL parallel the acquisition of language by children. Previous research has found that a consistent and effective use of spatial contrast develops relatively late in native-signing children's acquisition of mature sign languages (Schick, 1990). Children acquiring ASL are able to express the perspective of another character starting from age three, but initially do so using direct quotation and constructed action within embodied representations, in which the signer's body represents the body of the referent character. At about 5 years of age, native signers are able to establish and use consistent spatial locations for co-reference and verb agreement (Loew, 1984; Lillo-Martin, 1991). It is not until 7–10 years of age that signers fully master anaphoric and other "long-distance" uses of space, applying spatial loci consistently across a set of utterances to produce cohesive narratives (van Hoek et al., 1987, 1989; Bellugi et al., 1990; Emmorey, 2002). Though the use of space is clearly a fundamental aspect of mature sign languages, utilizing signing space to encode and maintain reference throughout a narrative is a complex and late-developing skill.

The development of the narrative skill required to express a narrator's perspective and that of multiple characters is similarly gradual and relatively late. Switching among these perspectives requires both cognitive maturity and linguistic skill, and children acquiring spoken language typically master it only in the middle-school years (Berman and Slobin, 1994). This protracted development may inform why, along with their overall less frequent use of referential shift devices, first-cohort signers frequently produce narratives situated from the perspective of a character, using first-person embodiment devices, rather than structuring the story from a narrator's perspective, even though the narrator perspective most closely resembles their own.

As we consider these parallels between sign language acquisition and emergence, it appears that a primary, or more basic representation of perspective is generated within an embodied, viewer-space format, with spatial contrasts along the front-to-back axis. Yet most, if not all, mature sign languages actively use a diagrammatic format, and use contrasts along the left-to-right axis. These conventions clearly are taking hold in NSL; indeed, by the second cohort, the left-to-right axis has become the preferred one for spatial contrasts. Why might a language change in this way? If we may speculate, the left-to-right axis might offer advantages in perceptual salience that enable signers to better

exploit the three-dimensional signing space. Signing space is used for a variety of grammatical functions, such as verb agreement, anaphora, and other types of co-indexation, which utilize non-manual as well as manual sign elements to identify particular locations near the signer's body. Signers can associate locations with particular referents, and then use pointing, eye gaze, and subtle movements of the head or torso relative to those locations to refer back to those referents (Thompson et al., 2006). Discriminating between less overt markers such as eye gaze and body movements may be easier with contrastive locations along the left-to-right axis than locations along the front-to-back axis. In other words, a glance to the left may be easier to discriminate from neutral eye-gaze than a glance forward. Moreover, the physical signing space is wider left-to-right than it is long front-to-back, allowing for a greater number of distinct locations. Since a signer cannot easily refer to locations behind the back, the use of the front-to-back axis realistically offers only one location in contrast to the signer's body. Thus, the left-to-right axis allows for more contrastive locations that are more easily distinguished.

Despite these advantages, there is a cost in adopting the left-to-right axis for spatial contrasts. Because movements to the left and right are symmetrical, they may be more difficult to encode or remember. Research in spatial cognition has found that people can differentiate and recall contrasts along asymmetrical axes, such as up vs. down and front vs. back, better than symmetrical ones like left vs. right (Franklin and Tversky, 1990; Bryant et al., 1992). Furthermore, the contrast between left and right depends on perspective. When talking about the location of non-jointly viewed locations in physical space, the signer's right can correspond to the listener's (or a character's) left. This ambiguity may explain why it takes time for a community to converge on conventions that use symmetrical relations for contrasts in reference.

Our examination of the emergence of referential shift devices in NSL has revealed that grammatical conventions for indicating shifts in perspective emerged over two sequential age cohorts of signers, who learned the language in its first two decades. The first cohort had a fair amount of variability in their production, but even so, their narratives already contained the seeds of lexical and spatial elements that would become more frequent, and possibly obligatory, in the language of the second cohort. Spatial devices appear to have emerged more slowly, but have recently become as prevalent as non-spatial, lexical devices. Previous work on NSL shows that second- but not first-cohort signers use consistent spatial language for other functions, and that the use of space to systematically assign semantic roles to the arguments of verbs emerged only with the second cohort. Spatial referential shift devices may have emerged later because they depend on the establishment of fundamental spatial conventions in the language. We conjecture that the systematic use of spatial devices in more local environments, such as within phrases and sentences, allowed them to be repurposed at the discourse level. Thus, while the modality of sign languages can ultimately engender the syntactic use of three-dimensional space, we propose that a language must first develop consistent and systematic local spatial contrasts before harnessing space for long-distance, abstract grammatical functions. The consistent use of spatial language and

the grammatical use of space for shifting reference did not spring up unaided in first-cohort signers. Rather, the second cohort of signers, as children, built upon the achievements of the first. In this way, two sequential age cohorts of children transformed Nicaraguan signing from its gestural seeds to the full, complex language it is today.

REFERENCES

Aarons, D., and Morgan, R. Z. (2003). Classifier predicates and the creation of multiple perspectives in South African Sign Language. *Sign Lang. Stud.* 3, 125–156. doi: 10.1353/sls.2003.0001

Bellugi, U., Lillo-Martin, D., O'Grady, L., and van Hoek, K. (1990). The development of spatialized syntactic mechanisms in American Sign Language. *SLR* 87, 183–189.

Berman, R. A., and Slobin, D. A. (1994). *Relating Events in Narrative: A Crosslinguistic Developmental Study*. Hillsdale, NJ: Lawrence Erlbaum.

Blyth, C., Recktenwald, S., and Wang, J. (1990). I'm like, "say what?!": a new quotative in American oral narrative. *Am. Speech* 65, 215–227. doi: 10.2307/455910

Bosworth, R. G., and Emmorey, K. (2010). Effects of iconicity and semantic relatedness on lexical access in American Sign Language. *J. Exp. Psychol. Learn. Mem. Cogn.* 36, 1573–1581. doi: 10.1037/a0020934

Bryant, D. J., Tversky, B., and Franklin, N. (1992). Internal and external spatial frameworks for representing described scenes. *J. Mem. Lang.* 31, 74–98. doi: 10.1016/0749-596X(92)90006-J

Chafe, W. (1982). "Integration and involvement in speaking, writing and oral literature," in *Spoken and Written Language: Exploring Orality and Literacy*, ed D. Tannen (Norwood, NJ: Ablex), 35–53.

Clark, H. H., and Gerrig, R. J. (1990). Quotations as demonstrations. *Language* 66, 764–805. doi: 10.2307/414729

Cormier, K., Smith, S., and Zwets, M. (2013). Framing constructed action in British sign language narratives. *J. Pragmat.* 55, 119–139. doi: 10.1016/j.pragma.2013.06.002

Coppola, M. (2002). *The Emergence of Grammatical Categories in Home Sign: Evidence From Family-Based Gesture Systems in Nicaragua*. Doctoral dissertation, University of Rochester, Rochester, NY.

Coppola, M., Gagne, D., and Senghas, A. (2013). "WHO chased the bird? Narrative cohesion in an emerging language," in *Linguistics Society of America Annual Meeting Extended Abstract* (Boston, MA).

Coppola, M., and Senghas, A. (2010). "Deixis in an emerging sign language," in *Sign Languages: A Cambridge Language Survey*, ed D. Brentari (Cambridge: Cambridge University Press), 543–569.

Emmorey, K. (1996). "The confluence of space and language in signed languages," in *Language and Space*, eds P. Bloom, M. Peterson, L. Nadel, and M. Garrett (Cambridge, MA: MIT Press), 171–209.

Emmorey, K. (2002). *Language, Cognition, and the Brain: Insights From Sign Language Research*. Mahwah, NJ: Lawrence Erlbaum.

Emmorey, K., and Falgier, B. (1999). "Talking about Space with Space: Describing environments in ASL," in *Storytelling and Conversation: Discourse in Deaf Communities*, ed E. A. Winston (Washington, DC: Gallaudet University Press), 3–26.

Emmorey, K., and Reilly, J. (1995). *Language, Gesture, and Space*. Hillsdale, NJ: Lawrence Erlbaum.

Engberg-Pedersen, E. (1993). *Space in Danish Sign Language: the Semantics and Morphosyntax of the Use of Space in a Visual Language*. Hamburg: Signum.

Engberg-Pedersen, E. (1995). "Point of view expressed through shifters," in *Language, Gesture, and Space*, eds K. Emmorey and J. S. Reilly (Hillsdale, NJ: Lawrence Erlbaum), 133–154.

Franklin, N., and Tversky, B. (1990). Searching imagined environments. *J. Exp. Psychol.* 119, 63–67. doi: 10.1037/0096-3445.119.1.63

Gernsbacher, M. A. (1997). "Coherence cues mapping during comprehension," in *Processing Interclausal Relationships: Studies in the Production and Comprehension of Text*, eds J. Costermans and M. Fayol (Hillsdale, NJ: Lawrence Erlbaum), 3–21.

Givón, T. (1995). "Coherence in text vs. coherence in mind," in *Coherence in Spontaneous Text*, eds M. A. Gernsbacher and T. Givón (Amsterdam: Benjamins), 59–116.

Goldin-Meadow, S., and Mylander, C. (1990). Beyond the input given: the child's role in the acquisition of language. *Language* 66, 323–355. doi: 10.2307/414890

Janzen, T. (2004). Space rotation, perspective shift, and verb morphology in ASL. *Cogn. Linguist.* 15, 149–174. doi: 10.1515/cogl.2004.006

Kegl, J., and Iwata, G. (1989). "Lenguaje de Signos Nicaraguense: a Pidgin sheds light on the "creole?" ASL," in *Proceedings of the Fourth Annual Meeting of the Pacific Linguistics Conference* (Eugene: University of Oregon).

Klima, E., and Bellugi, U. (1979). *The Signs of Language*. Cambridge, MA: Harvard University Press.

Labov, W. (1972). *Language in the Inner City: Studies in the Black English Vernacular*. Oxford: University of Pennsylvania Press.

Liddell, S. K. (1990). "Structures for representing handshape and local movement at the phonemic level," in *Theoretical Issues in Sign Language Research, Vol. 1: Linguistics*, eds S. D. Fischer and P. Siple (Chicago, IL: University of Chicago Press), 37–65.

Liddell, S. K. (2003). *Grammar, Gesture and Meaning in American Sign Language*. Cambridge: Cambridge University Press.

Liddell, S. K., and Metzger, M. (1998). Gesture in sign language discourse. *J. Pragmat.* 30, 657–697. doi: 10.1016/S0378-2166(98)00061-7

Lillo-Martin, D. (1991). *Universal Grammar and American Sign Language: Setting the Null Argument Parameters*. Dordrecht: Kluwer Academic Publishers.

Lillo-Martin, D., and Klima, E. S. (1990). "Pointing out differences: ASL pronouns in syntactic theory," in *Theoretical Issues in Sign Language Research, Vol. 1: Linguistics*, eds S. D. Fischer and P. Siple (Chicago, IL: University of Chicago Press), 191–210.

Lillo-Martin, D., and Meier, R. P. (2011). On the linguistic status of 'agreement' in sign languages. *Theor. Linguist.* 37, 95–141. doi: 10.1515/thli.2011.009

Loew, R. (1984). *Roles and Reference in American Sign Language: a Developmental Perspective*. Minneapolis, MN: University of Minnesota Press.

Meier, R. P. (1987). Elicited imitation of verb agreement in American Sign Language: iconically or morphologically determined? *J. Mem. Lang.* 26, 362–376. doi: 10.1016/0749-596X(87)90119-7

Meier, R. P. (1990). "Person deixis in American Sign Language," in *Theoretical Issues in Sign Language Research, Vol. 1: Linguistics*, eds S. Fischer and P. Siple (Chicago, IL: University of Chicago Press), 175–190.

Meir, I. (2012). The evolution of verb classes and verb agreement in sign languages. *Theor. Linguist.* 38, 145–152. doi: 10.1515/tl-2012-0008

Meir, I., Padden, C., Aronoff, M., and Sandler, W. (2007). Body as Subject. *J. Linguistics* 43, 531–563.

Meir, I., Sandler, W., Padden, C., and Aronoff, M. (2010). "Emerging sign languages," in *Oxford Handbook of Deaf Studies, Language, and Education*, Vol. 2, eds M. Marschark and P. Spencer (Oxford: Oxford University Press), 267–280.

Morgan, G., Herman, R., Barriere, I., and Woll, B. (2008). The onset and mastery of spatial language in children acquiring British Sign Language. *Cogn. Dev.* 23, 1–9. doi: 10.1016/j.cogdev.2007.09.003

Ochs, E. (1979). "Planned and unplanned discourse," in *Discourse and Syntax*, ed T. Givón (New York, NY: Academic Press), 51–80.

Padden, C. (1983). *Interaction of Morphology and Syntax in American Sign Language*. Unpublished doctoral dissertation, University of California, San Diego, San Diego, CA.

Padden, C. (1986). "Verbs and role shifting in American Sign Language," in *Proceedings of the Fourth National Symposium on Sign Language Research and Teaching*, ed C. Padden (Silver Spring, MD: NAD), 45–47.

Padden, C. (1990). "The relation between space and grammar in ASL verb morphology," in *Sign Language Research: Theoretical Issues*, ed C. Lucas (Washington, DC: Gallaudet University Press), 118–132.

Perniss, P. M. (2007). Achieving spatial coherence in German Sign Language narratives: the use of classifiers and perspective. *Lingua* 117, 1215–1338. doi: 10.1016/j.lingua.2005.06.013

Polich, L. (2005). *The Emergence of the Deaf Community in Nicaragua: With Sign Language You Can Learn So Much*. Washington, DC: Gallaudet University Press.

Poulin, C., and Miller, C. (1995). "On narrative discourse and point of view in Quebec Sign Language," in *Language, Gesture, and Space*, eds K. Emmorey and J. Reilly (Hillsdale, NJ: Lawrence Erlbaum), 117–131.

Pyers, J., Perniss, P., and Emmorey, K. (2008). "Viewpoint in the visual-spatial modality," in *Paper Presented at the 30th Annual Convention of the German*

Society of Linguistics Workshop on Gestures: a Comparison of Signed and Spoken Languages (Bamberg).

Pyers, J., and Senghas, A. (2007). "Referential shift in Nicaraguan Sign Language: a comparison with American Sign Language," in *Visible Variation: Comparative Studies on Sign Language Structure*, eds R. Pfau, P. Perniss, and M. Steinbeck (Amsterdam: Mouton de Gueyer), 279–302.

Pyers, J. E., and Senghas, A. (2009). Language promotes false-belief understanding: evidence from learners of a new sign language. *Psychol. Sci.* 20, 805–812. doi: 10.1111/j.1467-9280.2009.02377.x

Pyers, J. E., Shusterman, A., Senghas, A., Spelke, E. S., and Emmorey, K. (2010). Evidence from an emerging sign language reveals that language supports spatial cognition. *Proc. Natl. Acad. Sci. U.S.A.* 107, 12116–12120. doi: 10.1073/pnas.0914044107

Sandler, W., Meir, I., Padden, C., and Aronoff, M. (2005). The emergence of grammar: systematic structure in a new language. *Proc. Natl. Acad. Sci. U.S.A.* 102, 2661–2665. doi: 10.1073/pnas.0405448102

Schick, B. (1990). "The effects of morphosyntactic structure on the acquisition of classifier predicates in ASL," in *Sign Language Research: Theoretical Issues*, 358–374.

Schiffrin, D. (1981). Tense variation in narrative. *Language* 57, 45–62. doi: 10.1353/lan.1981.0011

Senghas, A. (1995). *Children's Contribution to the Birth of Nicaraguan Sign Language.* Unpublished doctoral dissertation, MIT, Cambridge, MA.

Senghas, A., and Coppola, M. (2001). Children creating language: how Nicaraguan Sign Language acquired a spatial grammar. *Psychol. Sci.* 12, 323–328. doi: 10.1111/1467-9280.00359

Shepard-Kegl, J. (1985). *Locative Relations in American Sign Language Word Formation, Syntax, and Discourse.* Unpublished doctoral dissertation, MIT, Cambridge, MA.

Supalla, T. R. (1982). *Structure and Acquisition of Verbs of Motion and Location in American Sign Language.* Unpublished doctoral dissertation, University of California, San Diego, San Diego, CA.

Slobin, D. I., Hoiting, N., Kuntze, M., Lindert, R., Weinberg, A., Pyers, J., et al. (2003). "A cognitive/functional perspective on the acquisition of "classifiers.","" in *Perspectives on Classifier Constructions in Sign Languages*, ed K. Emmorey (Mahwah, NJ: Lawrence Earlbaum), 271–296.

Tannen, D. (1982). Oral and literature strategies in spoken and written narrative. *Language* 58, 1–21. doi: 10.2307/413530

Tannen, D. (1986). "Introducing constructed dialogue in Greek and American conversational and literary narrative," in *Direct and Indirect Speech*, ed F. Coulmas (Berlin: Mouton), 11–32.

Taub, S. F. (2001). *Language From the Body: Iconicity and Metaphor in American Sign Language.* Cambridge, UK: Cambridge University Press.

Thompson, R., Emmorey, K., and Kluender, R. (2006). The relationship between eye gaze and verb agreement in American Sign Language: an eye-tracking study. *Nat. Lang. Linguist. Theory* 24, 571–604. doi: 10.1007/s11049-005-1829-y

van Hoek, K. (1992). Conceptual spaces and pronominal reference in American Sign Language. *Nordic J. Linguist.* 15, 183–199. doi: 10.1017/S0332586500002596

van Hoek, K., Norman, F., and O'Grady-Batch, L. (1989). "Development of spatial and non-spatial referential cohesion," in *Unpublished Paper Presented at Stanford Child Research Forum* (Stanford, CA).

van Hoek, K., O'Grady, L., and Bellugi, U. (1987). "Morphological innovation in the acquisition of American Sign Language," in *Papers and Reports on Child Language Development*, ed E. Clark (Stanford, CA: CSLI Publications), 116–123.

Winston, E. (1991). Spatial referencing and cohesion in an American Sign Language text. *Sign Lang. Stud.* 73, 397–410. doi: 10.1353/sls.1991.0003

The emergence of embedded structure: insights from Kafr Qasem Sign Language

Itamar Kastner[1], Irit Meir[2]*, Wendy Sandler[3] and Svetlana Dachkovsky[3]

[1] Department of Linguistics, New York University, New York, NY, USA
[2] Department of Hebrew Language and Department of Communication Disorders, University of Haifa, Haifa, Israel
[3] Department of English Language and Literature, University of Haifa, Haifa, Israel

This paper introduces data from Kafr Qasem Sign Language (KQSL), an as-yet undescribed sign language, and identifies the earliest indications of embedding in this young language. Using semantic and prosodic criteria, we identify predicates that form a constituent with a noun, functionally modifying it. We analyze these structures as instances of embedded predicates, exhibiting what can be regarded as very early stages in the development of subordinate constructions, and argue that these structures may bear directly on questions about the development of embedding and subordination in language in general. Deutscher (2009) argues persuasively that nominalization of a verb is the first step—and the crucial step—toward syntactic embedding. It has also been suggested that prosodic marking may precede syntactic marking of embedding (Mithun, 2009). However, the relevant data from the stage at which embedding first emerges have not previously been available. KQSL might be the missing piece of the puzzle: a language in which a noun can be modified by an additional predicate, forming a proposition within a proposition, sustained entirely by prosodic means.

Keywords: sign language, prosody, embedding, syntax, nominalization

INTRODUCTION

If a woman sits on a sofa and a man shows her a picture, can we say that *the man is showing the seated woman a picture*? We can in English and in many spoken languages; the participle *seated* allows us to express a secondary predicate which is subordinate to the main clause. As we started to investigate an as-yet undocumented young sign language from the town of Kafr Qasem in Israel we noticed an unexpected moderate tendency to use secondary predicates as noun modifiers. We regard this phenomenon as embedding, and situate it within two contexts. The first context is the overall emergence of structure in Kafr Qasem Sign Language (KQSL). The second is the question of how embedding and subordination may have evolved in natural language in general, over the ages. The latter is still a mystery for the most part, since we do not have documentation of early enough stages in the life of a language. The rise of embedding in KQSL, caught at a relatively early stage of development, could provide a clue to the initial stages of this phenomenon, and shed some light on the rise of syntactic complexity in general.

Since this is the first published work on KQSL, we begin by introducing the language, focusing on relevant historical and social aspects. We report on a brief study verifying that KQSL is not related to other sign languages in the region. We then describe our methods. The following sentence presents the results and analysis of the structures which we regard as embedded. The embedding findings arose while eliciting sentences for an investigation of word order, and we touch on this to put our main finding in perspective. To put this issue into a historical perspective, in Section Discussion: Embedding by Hand and by Mouth we compare possible origins of embedding that have been proposed for spoken languages with the KQSL findings. We summarize our findings and their theoretical relevance in the conclusion.

KAFR QASEM SIGN LANGUAGE

Sign languages arise spontaneously in communities of deaf people, and are not related to (though are possibly influenced by) the ambient spoken languages. In some countries, signers have by and large converged upon a single sign language, used by the deaf population throughout the country. Thus, deaf people in America use American Sign Language (ASL), while deaf people in Britain use British Sign Language (BSL), two mutually unintelligible languages. In Israel, the established language of the deaf community of about 10,000 signers is Israeli Sign Language (ISL). Yet Israel is home to a number of other sign languages which have arisen there over the past century, languages used by smaller, sometimes insular communities with an unusually high percentage of congenital deafness. Such languages are termed *village sign languages* (Meir et al., 2010; Nyst, 2012; Zeshan and de Vos, 2012) present other terms used in the literature to refer to these communities). Two village sign languages have already been documented in Israel. In the Bedouin village of Al-Sayyid, a community of about 4000 members of whom 130 are deaf, a sign language arose about 80 years ago (Al-Sayyid Bedouin Sign Language, ABSL), and, for more than a decade, has been the focus of intensive anthropological (Kisch, 2000, 2007, 2012) and linguistic investigation (e.g., Sandler et al., 2005; Aronoff et al., 2008; Meir et al., 2013; Sandler et al., in press). In the Jewish community of the sub-Saharan town of Ghardaia, Algeria,

another sign language emerged (Algerian Jewish Sign Language, AJSL). When the community members left Algeria and immigrated to Israel and France, they brought the sign language with them and continue to use it to this day (Lanesman and Meir, 2012a,b). To the list of sign languages being investigated in Israel, we now add KQSL.

The town of Kafr Qasem lies in the so-called Triangle area of Arab towns in central Israel and has existed for 350 years. Of its 20,000 residents, approximately 100 are deaf. From reports and interviews with residents of the town we estimate that the language is four generations old. Our team has been in touch with the local deaf community since early 2010, gathering social and historical data and analyzing the language's phonology, lexicon and syntax. This work has led us to conclude that KQSL is an independent language, worthy of study both for its own sake and in comparison to other languages. We give the results of a lexico-statistical study supporting this view, after providing some historical context.

HISTORY AND SOCIOLINGUISTIC CONTEXT

From what we have been able to uncover, an 80-year-old woman is the oldest living signer of KQSL. She is known to have had deaf aunts and uncles, placing our estimate for the age of the language at just under 100 years. Until the 1960s, the number of deaf people in the village was about 12. A rapid increase in the general population resulted in an increase in the number of deaf individuals, numbering about 30 by 1980, and over 100 today (Meyad Sarsur, pers. commun. 2013).

We do not possess detailed records of the deaf population over the last century. The history of the local community, gathered through interviews with people in the community, is as follows. A deaf woman from the south of the country married a hearing man from the village over 100 years ago, later giving birth to a number of deaf children. In the 1920s, 1930s and 1940s there were 1800 inhabitants in the village, out of whom 12 were deaf (7 male, 5 female). By the early 1980s the number reached 31 (14 male, 17 female). The rise in the number of deaf children led to the opening of a class for deaf children in the local school in 1979. Although the first teacher was a non-signer, in 1985 a teacher who used signing started working in school, introducing ISL vocabulary into the educational system there. Since 1993, cochlear implants have been available through the Israeli health system; around 30 children have been implanted. According to one of our deaf consultants, the parents of these children "reject the use of sign language." A number of parents of deaf children founded a local association in 1995, paving the way for a deaf club for afternoon activities with sign language which opened in 1996. Today the town has a number of educational programs available for deaf children: a local branch of *MICHA*, the Israeli preschool system for young deaf children; Learning Centers; a kindergarten; elementary school classes in the nearby town of Jaljulia; and classes for deaf and hard-of-hearing children in the local junior high and high school.

From our interviews with deaf and hearing people who are 60 years and older, we have learned that the older deaf people in the village spent a lot of time with each other as a group in their childhood. They have a sense of "togetherness" that has persisted to this day. There are many social meetings held at the house of one of the older women which serves as the local gathering place for deaf people of all ages, situated on a street nicknamed "the deaf neighborhood." Some of the deaf people are married, usually to a hearing spouse, though there is at least one deaf-deaf marriage in the younger generation. The hearing spouses, siblings, children and neighbors of deaf people communicate with them in sign. Hearing people that we interviewed report that this has been the practice for as long as they remember, which goes back about 60–70 years. Since contact with the Jewish deaf community and the general educational system for the deaf in Israel began around the late 1970s, we assume that the sign language that emerged in the village up to that point developed independently of ISL. We have found no evidence of contact with the better studied village sign language in Israel, ABSL, over the years. This is not surprising, due to geographical and cultural distances between the two. Even today, contact between people from the two communities is very rare. We therefore conclude that KQSL developed as an independent local sign language. This conclusion is corroborated by the lexical comparative study reported below.

However, deaf people in Kafr Qasem who entered the educational system in the 1980s and onwards have been heavily influenced by ISL, and many of them find it hard to understand the older people who use only KQSL. In order to document and describe the linguistic structure of KQSL, it is necessary to study the language of those signers who have had little contact with ISL over the years, and who use KQSL regularly as their main means of communication. The study reported here is based on the sign production of six such KQSL signers.

EVIDENCE FOR KQSL AS AN INDEPENDENT LANGUAGE

One clear indication that two languages are unrelated is the existence of two different lexicons. In fact, one of our first impressions of KQSL was that its vocabulary is unlike that of ISL or ABSL, and we therefore had to be aided by an interpreter to interact with signers there. In the KQSL lexicon, we have found signs for tangible objects, abstract concepts, actions and feelings. Many of these are directly related to the local culture. For instance, KQSL has two signs for the concept *sheep*, each referring to a different type of sheep, a distinction important for a community that used to engage in herding. The lexicon also includes signs that can be regarded as function words, such as negators, signs denoting quantity and signs denoting degree.

In order to determine the degree of overlap between KQSL and the neighboring sign languages, we follow previous work in comparing the lexicons of sign languages and their dialects (McKee and Kennedy, 2000; Guerra Currie et al., 2002; Hendriks, 2008; Al-Fityani and Padden, 2010). The methodology is based on comparing the relative resemblance between signs in different languages that denote the same concept. To this end we require definitions of what *identical* signs and what *similar* signs are, and so a word on sign language phonology is in order.

Contrastive features in established sign languages are hierarchically organized into a number of major phonological categories: handshape, location and movement (see Sandler, 2012 for an overview). Although there is some debate regarding the status of palm orientation in the system, i.e., whether it is a fourth

major category or subordinate to handshape (Sandler and Lillo-Martin, 2006, p. 156), this theoretical point is not of great import here; features of orientation are clearly distinctive in sign languages, and this suffices for our purposes. That these categories contain contrastive features can be readily seen through examination of minimal pairs, using ISL as an example. In ISL, the signs DEPRIVED and PROFIT (**Figure 1A**) are distinguished by features of the two handshapes and . This is a minimal pair, because the locations and movements are the same in the two signs, which are distinguished by handshape alone. The ISL signs WELL-BEING and CURIOSITY (**Figure 1B**) are minimally distinguished by features of location (chest vs. nose respectively), while ESCAPE and BETRAY are distinguished by movement alone, straight for ESCAPE, and arc for BETRAY (**Figure 1C**).
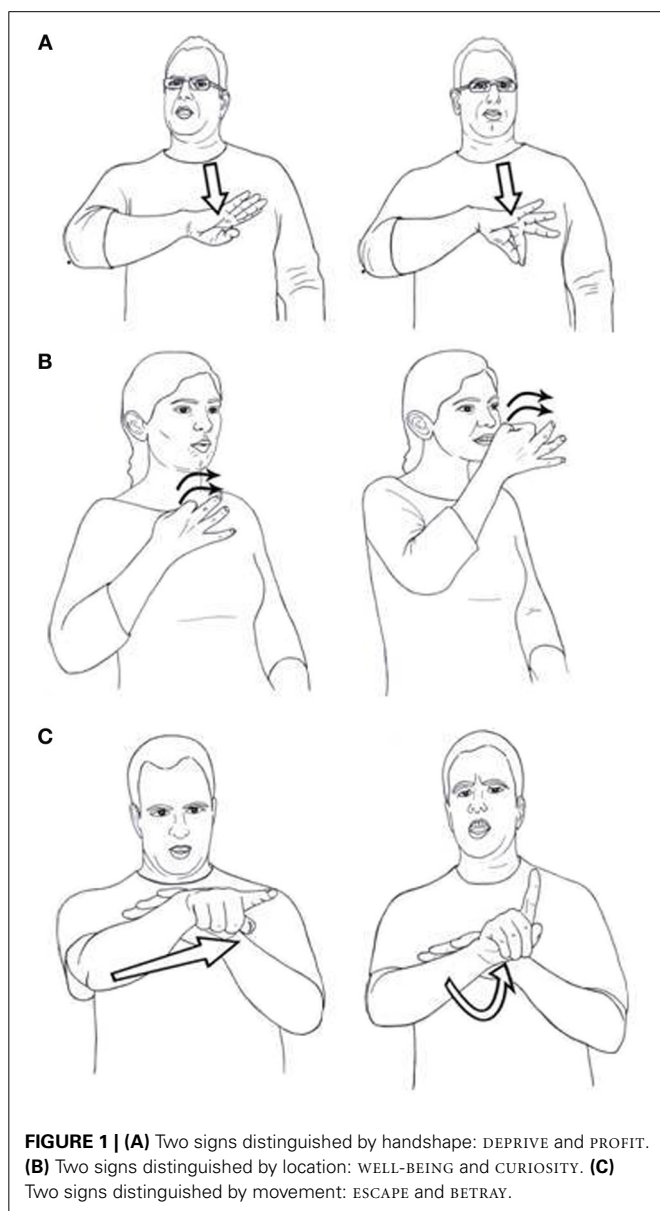
These phonological features are used when comparing the lexicons of different sign languages in order to determine the degree of overlap between them. Following McKee and Kennedy (2000, p. 51), we define *identical* as signs which are pronounced with the same handshape, location, movement and orientation. *Similar* signs differ in only one of these parameters. Comparing ISL with KQSL signs, we find a few signs in the two languages that are identical; they share all four components, as is illustrated by the sign for "bird" in **Figure 2A**. The signs for "television" (**Figure 2B**) are similar; they have the same handshape, location and movement, but differ in orientation. The signs for "cow" (**Figure 2**C) are different; they differ on more than one phonological component. They have a different handshape and a different movement.

Even sign languages that we know to be genetically unrelated might display substantial similarities in vocabulary. Previous work has shown that unrelated sign languages normally display between 20 and 30% overlap in their lexicons (McKee and Kennedy, 2000). For example, ASL shares 26.3 and 24.5% identical signs with two languages which are genetically related to each other but not to ASL itself, namely BSL and Australian Sign Language, respectively. Including similar signs in the tally raises the rates to 32.6 and 32.7% (McKee and Kennedy, 2000, p. 53).

Guerra Currie et al. (2002, p. 228) obtained similarity ratings of 38% in 112 sign pairs when comparing the distantly related Mexican Sign Language and French Sign Language; 33%

FIGURE 1 | (A) Two signs distinguished by handshape: DEPRIVE and PROFIT. **(B)** Two signs distinguished by location: WELL-BEING and CURIOSITY. **(C)** Two signs distinguished by movement: ESCAPE and BETRAY.

FIGURE 2 | (A) Two identical signs in ISL and KQSL: bird. **(B)** Two similar signs in ISL and KQSL: TELEVISION. **(C)** Two different signs in ISL and KQSL: COW.
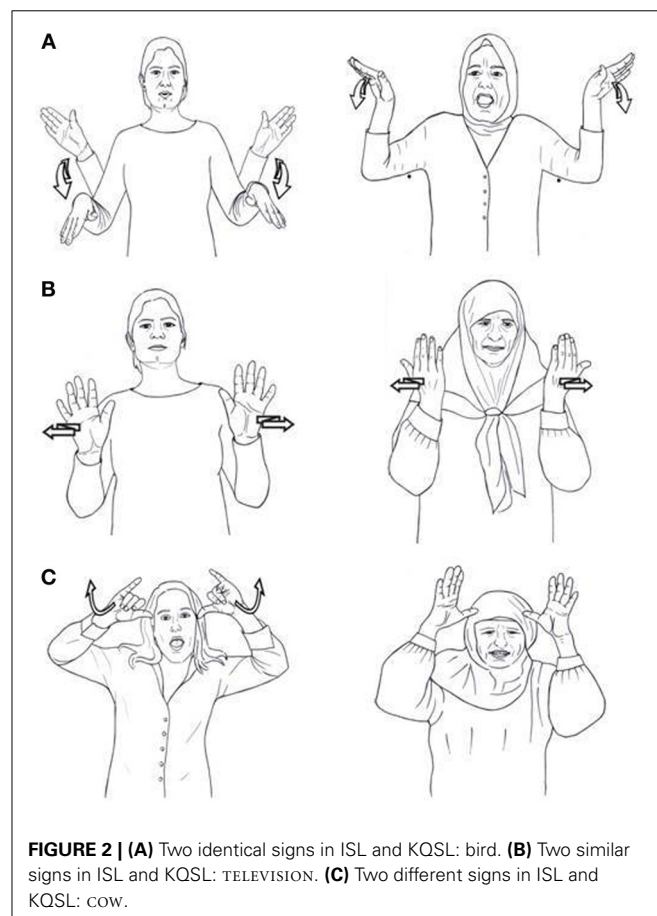
similarity in 89 pairs between the culturally linked, but not genetically related, Mexican Sign Language and Spanish Sign Language; and 23% similarity in 166 pairs between the unrelated Mexican Sign Language and Japanese Sign Language.

This state of affairs, in which unrelated sign languages show such similarities, is attributed to iconicity, pervasive in the sign language lexicon. Two signs denoting the same concept in two different languages may represent iconically the same aspect of the concept being described. In such cases, they will display similarity in form, whether or not the two languages are related. For example, a sign for "eat" might represent food going into the mouth. In this sign, the location of the sign will be the mouth and the hand will move toward the mouth. The shape of the hand might vary from language to language, as it does in ISL and ABSL for example, shown in **Figure 3**. According to McKee and Kennedy's criteria, these two signs are *similar*. However, this similarity does not necessarily reflect a genetic relationship between the two languages. Even when the features of all four parameters are identical in two sign languages, as they are for EAT in ISL and in ASL, one would not wish to use this coincidence as evidence for proximity on a sign language family tree. The two signs may be identical because they are built on the same mental image representing the concept. It is for this reason that a percentage as high as 30% similarity between the lexicons of two sign languages should not be interpreted as reflecting a family relationship. In contrast, British, New Zealand, and Australian Sign Languages are very closely related, with 82% identical signs from a Swadesh list, and 98% similar signs (McKee and Kennedy, 2000; Johnston, 2003).

Returning to our comparison of KQSL, ISL, and ABSL, we have been using an adapted Swadesh list of concepts for comparison of different dialects and signers (we added a number of concepts that are likely to exist in all sign languages of the region, such as "Jerusalem"; the list is given in Supplementary Materials). Using this list we conducted three pairwise comparisons of elicited citation forms between ISL, KQSL, and ABSL. Coding was done by the first author and checked by the second author and two deaf consultants. Cases of disagreement were discussed until a consensus was arrived at. We compared 161 pairs of signs that exist as lexical signs in both KQSL and ABSL, finding a 19% overlap in identical signs which rose to 36% when similar signs were

included. The overlap is similar when comparing KQSL and ISL: of the 186 pairs of signs that exist in both languages, 15% show overlap when considering identical signs and 36% overlap when including similar signs. The comparison of 161 pairs of signs in ABSL and ISL showed lesser degree of overlap: about 9% overlap with identical signs and 23% overlap when similar signs are included. The results of this comparison are given in **Figure 4**.

This suggests that, from a lexico-statistical point of view, KQSL and ISL are no more related than ASL and BSL are to one another (31% identical in the latter case). In spite of cultural similarities between the communities of KQSL and ABSL, their lexicons show a very similar degree of overlap. In some cases, similarity or identity between pairs of signs in these two languages may be partly due to some shared aspects of their culture, which are represented iconically by their vocabulary. For example, signs for "man" in KQSL and ABSL take the moustache to be the distinguishing feature of a man. This is a typical characteristic of men in Arab communities in the region, and is reflected in the choice of the mental image underlying the signs in both languages. In ISL, as in many European based sign languages including ASL, the sign for "man" is articulated on the forehead, maybe iconically representing a cap. The sign for "sheep" in ISL represents a wooly body, while the KQSL and ABSL signs represent the wobbly tail of the sheep, a very noticeable feature if you are a shepherd walking behind your herd (see **Figure 5**). Though the two signs are different, they might be regarded as similar since they differ only in orientation: in ABSL the fingertips are oriented downwards whereas in KQSL the fingertips point forward. Yet their similar form may reflect shared cultural practices rather than linguistic affiliation.

In spite of shared cultural practices and iconicity, the lexicons of KQSL and ABSL are different in about 65% of the items in our list, as are the lexicons of KQSL and ISL. From interviews with people in the Kafr Qasem and Al-Sayyid communities we learned that contact between members of the communities has been very rare and sporadic, and contact between deaf ISL signers

FIGURE 3 | EAT in ISL and ABSL.

FIGURE 4 | Comparison of lexical similarity between ISL, KQSL, and ABSL.
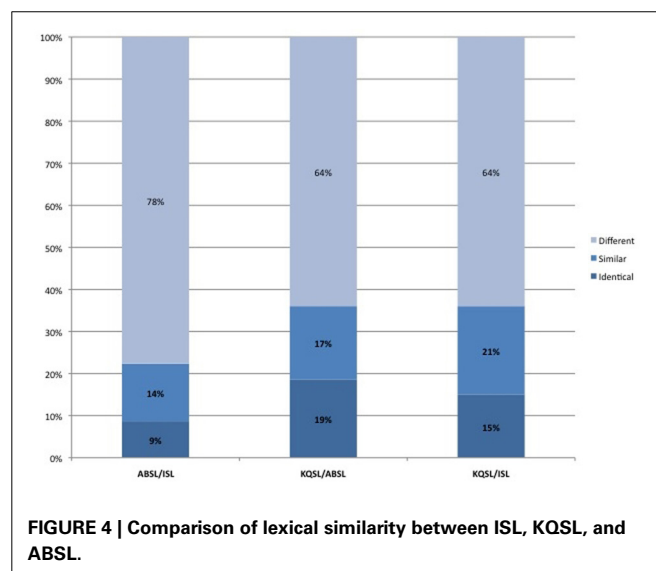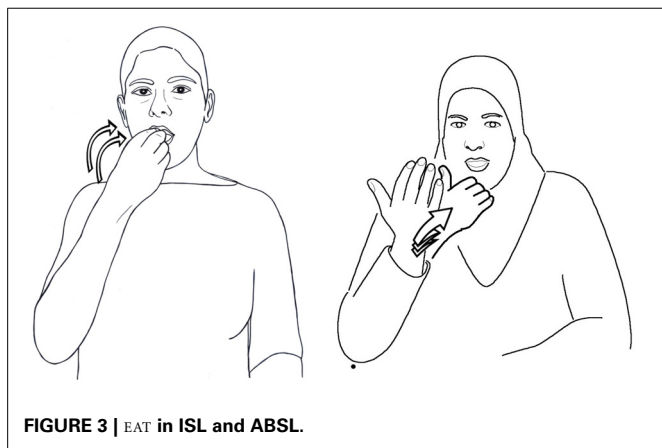
FIGURE 5 | The sign for sheep in ISL, KQSL, and ABSL.

and the older members of the KQSL deaf community has also been limited. Since the lexicons are different and the historical and social backgrounds are different, it seems reasonable to conclude that these are different languages with potentially different grammatical characteristics. Next we turn to the details of our study.

METHODS

PARTICIPANTS

Six native KQSL signers participated in the study, five deaf and one hearing. The participants were of the 2nd and 3rd generation of Kafr Qasem deaf, with ages ranging between 42 and 67 ($M = 52.7$), four female and two male. Our group included one father-daughter pair. The results come from two of the female participants and one of the male participants. All participants received an explanation in sign language (ISL, which was then translated to KQSL by a member of the community) about the goals of the research and the methodological procedure. Only those participants who gave their consent to participate were included in the study.

MATERIALS

We used 30 short clips originally compiled for Sandler et al. (2005). The elicitation material included 13 intransitive sentences, 13 monotransitive sentences and four ditransitive sentences (the list is given in Supplementary Materials). Of the monotransitive sentences, eight portrayed a person acting on an object (e.g., a man washing a plate) and five portrayed an interaction between two humans (e.g., a girl combing a woman's hair). All four ditransitive sentences had two animate participants (e.g., a woman handing a shirt to a man).

PROCEDURE

Participants viewed these video clips on a laptop, one clip at a time, and then described them in sign language to another native signer seated across from them. The interlocutor was asked to choose one of three pictures portraying the scene on a printed page, in order to verify comprehension. For example, if the video showed a woman giving a shirt to a man, the page included a picture of a woman giving a man a shirt, a picture of a man giving a woman a shirt and a picture of a woman looking at a man.

If an incorrect picture was chosen, the signer was asked to produce the sentence again. Some of the subjects signed the same sentence more than once of their own accord. Conditions were pseudo-randomized in advance, creating two lists with different item orders, so that each participant was shown the stimuli in one of two orders.

PROSODIC ANALYSIS

The elicited material underwent a preliminary gloss by a deaf research assistant, a native signer of ISL who has been in contact with KQSL signers for a number of years. Next, a quadrilingual consultant (KQSL, ISL, Arabic, and Hebrew) watched all elicited utterances and was recorded translating each one into Hebrew. This consultant, a trained ISL interpreter, is fluent in Arabic, KQSL, ISL, and Hebrew but was more comfortable volunteering simultaneous translations into Hebrew than sign-by-sign glosses. The preliminary glosses were compared to the translations provided, at which point the authors then discussed the best way to gloss each sign. Once agreement was reached, the gloss for each utterance was checked with the consultant once more to reach the final version reported here. Data were glossed and coded using the ELAN annotation system, which also provides the time windows for each annotation, allowing us to measure sign duration[1].

The elicitation and glossing procedures were designed to analyze basic clause structure in KQSL, particularly word order. However, the first step toward defining clause structure is to define clause boundaries. In a language that has not been previously studied, this is not a trivial matter, since there is no pre-existing information regarding properties of clauses in the language. A dilemma regarding parsing of a stretch of discourse into clauses immediately arises when two signs denoting actions occur in the same response, as in: MAN SIT GET-UP. How should this stretch of signs be analyzed? As a coordination of two clauses, e.g., "the man sat down and then stood up," or maybe as one clause containing a main predicate and a secondary

---

[1]The ELAN tool was developed at the Language Archive, Max Planck Institute for Psycholinguistics, Nijmegen (http://tla.mpi.nl/tools/tla-tools/elan/). See Crasborn and Sloetjes (2008) for a discussion of the use of ELAN to code sign language data.

predicate, as in "the seated man (the man who was sitting) stood up"? Since we do not know anything about clause structure in the language, the initial analysis cannot be based on syntactic properties. In such cases we found that prosodic cues are very helpful.

In previous analyses of clause structure in unstudied sign languages (ABSL, Sandler et al., 2005; Padden et al., 2010) and ISL (Meir, 2010), a method was developed for determining clause boundaries based on the semantics and the prosody of the signs. Semantically, a clause is a unit containing a predicate, and associated signs are determined by thematic roles associated with the predicate. Prosodic cues, such as shifts in the rhythm marked by a pause or lowering of the hands, together with a change in head or body position, determine the boundaries of an intonational phrase, which often corresponds to a clause (Nespor and Vogel, 1986). Furthermore, prosodic cues also mark constituents within a clause, which often correspond to smaller syntactic constituents such as phrases.

The prosodic analysis employed here relies on the model of sign language prosody developed in Sandler and colleagues' work mainly on ISL (Nespor and Sandler, 1999; Dachkovsky and Sandler, 2009; Sandler, 2011; Dachkovsky et al., 2013). The prosodic cues in the list below were used as indicators of constituent boundaries, in accord with criteria developed in earlier work. These cues include manual timing and certain non-manual markers. These cues are typically combined at major prosodic constituent boundaries, i.e., intonational phrases. Specifically, increasing the salience of the final sign in a constituent through lengthening, holding the hands in place, or repeating the sign is accompanied by and aligned with a concomitant shift in head position and change of facial expression (Dachkovsky and Sandler, 2009; Dachkovsky et al., 2013). Smaller constituent boundaries (e.g., phonological or intermediate phrase boundaries) may be similarly marked by changes in hand rhythm and facial expression, but typically not by change in head or body position. Duration is increased at final prosodic constituent boundaries (Nespor and Sandler, 1999). An exception is constituent-final pronouns, which can cliticize onto their preceding hosts, and be observably reduced in duration (Sandler, 1999). Similar criteria have been used for determining constituent boundaries in previous work on ABSL (Sandler et al., 2005, 2011; Padden et al., 2010). The earlier literature does not provide measures of the duration of signs other than the final signs within constituents, a measure which we have found to be useful for the phenomenon under discussion in KQSL. Of the cues reported in the literature, the following are relevant for the present study:

(1) Pause or hold. A pause in signing is defined as a period of transition during which the hands are relaxed. A hold is when the hands are held still at the final location for the sign (Liddell and Johnson, 1989). One or the other of these cues is typically found at the end of a prosodic phrase (Nespor and Sandler, 1999).

(2) Change in body position. Among other uses of torso movement, such as role shift in a discourse, movement of the upper body can indicate the start of a new constituent

(Boyes Braem, 1999 for Swiss German Sign Language, Fenlon et al., 2007 for Swedish Sign Language and BSL, Nicodemus, 2009 for ASL, Herrmann, 2009 for German Sign Language, Jantunen, 2007 for Finnish Sign Language, Van der Kooij et al., 2006 for Sign Language of the Netherlands, and see Ormel and Crasborn, 2012 for an overview).

(3) Change in head position. The head turns to one side or the other, or moves forward or back[2].

(4) Lengthened sign duration and/or increased size beyond those of citation form.

(5) Spread of facial expressions. While grammatical facial expressions, such as raised brows on yes/no questions, function as intonation (Nespor and Sandler, 1999) and typically characterize a whole intonational phrase (Liddell, 1980), some lexical signs are accompanied by a specific facial expression of their own. If this expression is spread to adjacent signs, we take this to be an indication that they belong to the same prosodic constituent (see discussion in Crasborn et al., 2008)[3].

RESULTS

Here we report the word order results briefly, and go on to provide a detailed analysis of the embedded structure we found in our data.

WORD ORDER

In total, 213 elicited utterances were recorded. Subject-Object-Verb (SOV) is the predominant order, occurring in over 63% of the responses, as in (1). OSV and SVO follow up with 19 and 14%, as in (2) and (3) respectively. We may conclude that SOV is the prevalent word order in KQSL. The remaining 4% are divided between less frequent orders such as SV, OV, and OVS (4); for a more detailed analysis of word order in KQSL compared to other languages see Meir et al. (in preparation)[4].

(1) SOV: WOMAN MOTHER MAN SEE (1.2.04b)
'A woman looks at a man'

(2) OSV: WOMAN MAN BOOK LOOK (1.3.14)
'A man shows a picture to a woman'

(3) SVO: FATHER FATHER WASH-DISHES PLATE (1.5.21)
'A man washes dishes'

(4) Other: GIRL LITTLE PUSH (1.2.03)
'A girl drags a shopping cart'

---

[2]For differences prosodic uses of head movement in ISL and ASL, see Dachkovsky et al. (2013).

[3]Unlike the other criteria that we used, which are all supported in detail by earlier research, facial expression spread from a specific word has not been employed in earlier work and requires further confirmation. However, the criterion is lent credence by its similarity to another previously supported phenomenon: the mouthing of spoken words in ISL, which spreads from the specific sign it accompanies to neighboring signs within small prosodic units (Sandler, 1999).

[4]The emergence of a prevalent word order does not mean that all clauses in KQSL contain a straightforward arrangement of subject, object and verb. See examples (16) and (17) below for utterances in which some arguments or predicates are repeated.

EMBEDDING

Identifying an embedded predicate

Out of the 213 responses, 10 presented us with a challenge: they contained two predicates (two signs denoting an action or a property), yet the two signs seemed to differ in both their function in the string and in their prosody. We begin with an in-depth analysis of a representative utterance. Consider the following stretch of signs, describing a clip in which a man and a woman are sitting on the sofa, and the woman is looking at the man:[5]

> (5) WOMAN MAN SIT EYE^LOOK-AT++.[6] (1.1.04)

The considerations we describe here were the same for each of the 10 examples analyzed. The first dilemma was whether the sign SIT is predicated of both the woman and the man ("the woman and the man are sitting"), or whether it is predicated only of the man. Prosodic cues provided the answer.

After the sign WOMAN, whose movement lasted about 400 ms, there was a hold in the movement of the hands (that is, the hands were held still at the final location of the sign) for about 250 ms. The end of the sign and the duration of the hold were aligned with a shift in both torso and head positions, the two bobbing slightly down and back up. This combination of features is typical of a prosodic boundary.

In contrast to this, there was no major postural change between MAN and SIT. More importantly, there was no hold of the hands or pause after the sign MAN; rather, the signer immediately transitioned to the beginning of the following sign, SIT. Instead, both the body posture and the behavior of the non-dominant hand linked MAN and SIT within a single prosodic constituent. First, a change in head and body posture after WOMAN characterized the two signs MAN SIT. The sign SIT is often accompanied by an upward movement of head and torso. In example (5), this upward posture of the torso starts on MAN, spreading regressively from SIT to MAN. Additionally, the non-dominant hand, which is not used in the sign MAN but comes into use in SIT, starts moving upwards toward the initial location of SIT while the dominant hand still signs MAN. This process of Non-dominant Hand Spread within (but, crucially, not beyond) phonological phrases has been described for ISL and compared with external sandhi phenomena that take place within prosodic constituents (Nespor and Vogel, 1986; Nespor and Sandler, 1999).

The overall effect is one of a smooth transition between MAN and SIT, as opposed to a marked break between WOMAN

---

[5]Following convention, we use small caps to denote glosses of individual signs. The caret ^ is used to denote compounding. Multiple plus signs ++ denote repetition. A hyphen indicates that two words in the gloss correspond to one sign. For example, the two words LOOK-AT correspond to one sign in KQSL. Full stops indicate clausal intonation break, and commas indicate minor intonational breaks, that is, intonational breaks within a clause. Square brackets [ ] delimit the prosodic constituents whose position and length are relevant for our analysis, as described in the text. Some signs have different versions (synonyms), and these are glossed as e.g. WOMAN1 WOMAN2. IX denotes an index (pointing in space). Each utterance is followed by its identifier in our database. Where relevant, sign duration in milliseconds is given as a subscript.
[6]The sign LOOK-AT is signed with a pointing sign towards the eye and then a V hand moving outwards. They form one lexical unit, "look-at."

and MAN. In other words, the prosodic pause between the sign WOMAN and the sign MAN, and the assimilatory hand and body movement between MAN and SIT, indicate that MAN SIT is a constituent, to the exclusion of WOMAN (in **Figure 7** below, the relevant head and body postures are indicated with dotted lines). It is thus unlikely that SIT modifies both WOMAN and MAN ("a woman and a man are sitting"), but very likely that it modifies MAN. In still illustrations, the effects we describe appear quite subtle, but the movement of the body in actual signing makes both spreading and changes of postures more salient.

Next we examined the prosody of the transition from SIT to the following sign, EYE^LOOK-AT. Here we see a break between the two signs, characterized by a change in head posture and body posture (head and body tilt to the left), dropping of the non-dominant hand, which was active in the sign SIT, before the sign EYE^LOOK-AT, and a change in rhythm: EYE^LOOK-AT is slower. However, there is no pause or hold between the two signs, making the boundary less prominent than an intonational phrase boundary which typically delineates clauses (Nespor and Sandler, 1999), and we therefore do not interpret it as corresponding to a clause boundary.

What then is the structure of this utterance? The first two prosodic constituents are the two arguments of the transitive event: WOMAN and MAN. The last constituent is the predicate (EYE^LOOK-AT). The order of constituents corresponds to the prevalent order in the language: it is SOV. However, we were baffled about how to analyze the verb SIT that formed a prosodic unit with MAN, and did not itself mark the end of the clause. We then noticed that SIT was much shorter in duration than the sentence-final verb, LOOK-AT, and much smaller in size. Let us make these measures explicit.

We use the term *shorter* signs to refer to the total duration (in milliseconds) of the sign production. We define the beginning of a sign as the point when the hands are fully in the handshape and orientation of that sign. The sign ends when they are no longer in that configuration. In the example above, SIT lasted 170 ms whereas LOOK-AT lasted 400 ms. In a subsequent utterance (6), the same sign, SIT, lasted 490 ms, showing that it is not the sign itself which is short but rather the specific production in sentence (5) above.

> (6) WOMAN SIT. ONE^GIRL SMALL BRUSH++. (1.1.22)
> 'A woman is sitting. A girl is brushing her hair.'

When we say *smaller* signs, we refer to the amount of physical space "taken up" by the sign. This measure is more difficult to quantify, yet when a sign occurs in two responses, the size of the sign in the two productions can be compared. We illustrate this with the verb SIT. In response (6), the two hands begin at chest level and are lowered to the hips, with each hand starting off with the palm faced outwards and the pinky finger near the ipsilateral shoulder (see **Figure 6A**). In example (5) above, however, the situation was slightly different. The dominant hand started off at chest level, while the non-dominant hand was a bit below it, such that the tips of the fingers in the non-dominant hand were in line with the palm of the dominant hand. The hands then moved slightly downwards, stopping just under the chest (see **Figure 6B**).

FIGURE 6 | (A) SIT (KQSL), non-small. **(B)** SIT (KQSL), small.

These two measures—shortened sign duration and reduced sign size—combine to give us a notion of phonological reduction. Taken together, the phonological reduction and the prosodic cues on and between WOMAN, MAN, SIT, and LOOK-AT indicate that in this production, SIT does not behave as a main predicate in a clause: it forms one constituent with the preceding sign MAN, to the exclusion of the sign WOMAN; there is a break between SIT and LOOK-AT, signaled by change in head and body posture and change in rhythm, but it is not a typical major boundary, as there is no hold or pause between the two signs; and SIT is much shorter than the other verbal sign in the clause, LOOK-AT. The full prosodic analysis of this clause is presented in **Figure 7** below[7]. "Duration" notes the total length of each sign in milliseconds, "Hold" indicates that the hands were held in position at the end of a sign, and "Big" means that the size of the sign takes up more physical space than the citation form. As pointed out above, we take the prosody to signal syntactic constituency as well, and we therefore analyze SIT as a secondary predicate in the clause, a modifier of MAN. An acceptable translation—confirmed by our consultant—would thus be "The woman is looking at the sitting (seated) man." We conclude, then, that the response in (5) is a clause containing two predicates, a main predicate and a secondary predicate. The two predicates differ in their position in the clause and in their form: the main predicate occurs in clause final position, and is longer and larger in form. The secondary predicate follows a nominal sign and forms a constituent with it and is reduced in both size and duration. The structure of the clause is as follows:

(7) [WOMAN] [[MAN SIT] [EYE^LOOK-AT]]
'The woman is looking at the sitting (seated) man'

Alternative hypotheses

Possible objections to the proposed analysis concern the difference in size and duration between the two predicates. It might be argued that LOOK-AT is longer than SIT because it is in utterance final position, because it is verbal rather than nominal or because it is reduplicated. Let us consider these objections one by one.

Signs occurring utterance-finally tend to be longer and bigger (Nespor and Sandler, 1999). Accordingly, SIT might be shorter not because it is embedded but rather because it is not in utterance final position. But since the prevalent word order in KQSL is SOV, there should be a strong tendency for the main predicate to occur in the clause final, prosodically prominent, position. How then can we show that the short duration of what we take to be an embedded predicate is indeed a signal of embedding rather than the result of its position in the utterance? Luckily, we found at least one example where the embedded predicate is in utterance final position while the main predicate is not:

(8) MOTHER. IX MOTHER WOMAN SIT. GIRL BRUSH-OTHER$_{450}$++ BRUSH$_{300}$++, [MOTHER SIT$_{210}$] (1.5.22)
'A mother is sitting. A girl is brushing the sitting mother's hair.'

BRUSH-OTHER and BRUSH are two variants of the "brushing" sign, with different hand orientations and locations; the former portrays a brushing action on an imaginary brushee and the latter on the signer herself. All sign durations are given for the non-reduplicated version, meaning one movement of the hands, even if the sign was reduplicated. For example, the first movement of BRUSH-OTHER in (8) took 450 ms, but in that utterance this movement was then repeated once more for a total duration of 800 ms for the entire, reduplicated sign. For the purpose of comparing main predicates and secondary predicates we calculated the duration of the main predicate without reduplication. In (8), the main verb, BRUSH-OTHER, is twice as long as the verb SIT (450 and 210 ms., respectively), though SIT is in final position and BRUSH-OTHER occurs in clause medial position. We conclude, then, that the relative duration and size of the predicate are indicative of its status as embedded or not.

The second objection concerns the status of grammatical categories in the language. It has been noted for various sign languages that nouns are reduced in size when compared to related verbs (e.g., Johnston, 2001 on Australian Sign Language; Hunger, 2006 on Austrian Sign Language; Kimmelman, 2009 on Russian Sign Language). It might be argued that SIT is shorter than LOOK-AT as it is a nominal in a modifier position and not a "real" verb. We do not know enough about the differences between nouns and verbs in KQSL to address this question. Yet examples (19) and (20) below suggest that we should be cautious in drawing such a conclusion. The sign HOUSE appears in both examples, yet in (20) it is three times longer than in (20). This difference shows that the part-of-speech status of a sign cannot be the only factor determining its duration; prosodic factors play a role too. In the absence of a detailed analysis of the properties of nouns and verbs in the language, we prefer to couch our analysis in prosodic terms, since such an analysis is not based on a syntactic analysis which we do not have. The relationship between prosodic factors and parts of speech deserves future attention as we learn more about the behavior of nouns and verbs in KQSL.

---

[7]In the videotape of the sentence illustrated in **Figure 7**, the dominant hand is beginning to transition from the handshape for MAN (fist) to the shape for SIT (flat, open hand), while still at the location for MAN, at the same time that the nondominant hand moves toward the position for SIT. The coarticulatory movement was so fast that only a blurred transitional shape was visible, and we elected to represent the handshape of the sign MAN in **Figure 7**, rather than the blurred transition, for clarity.
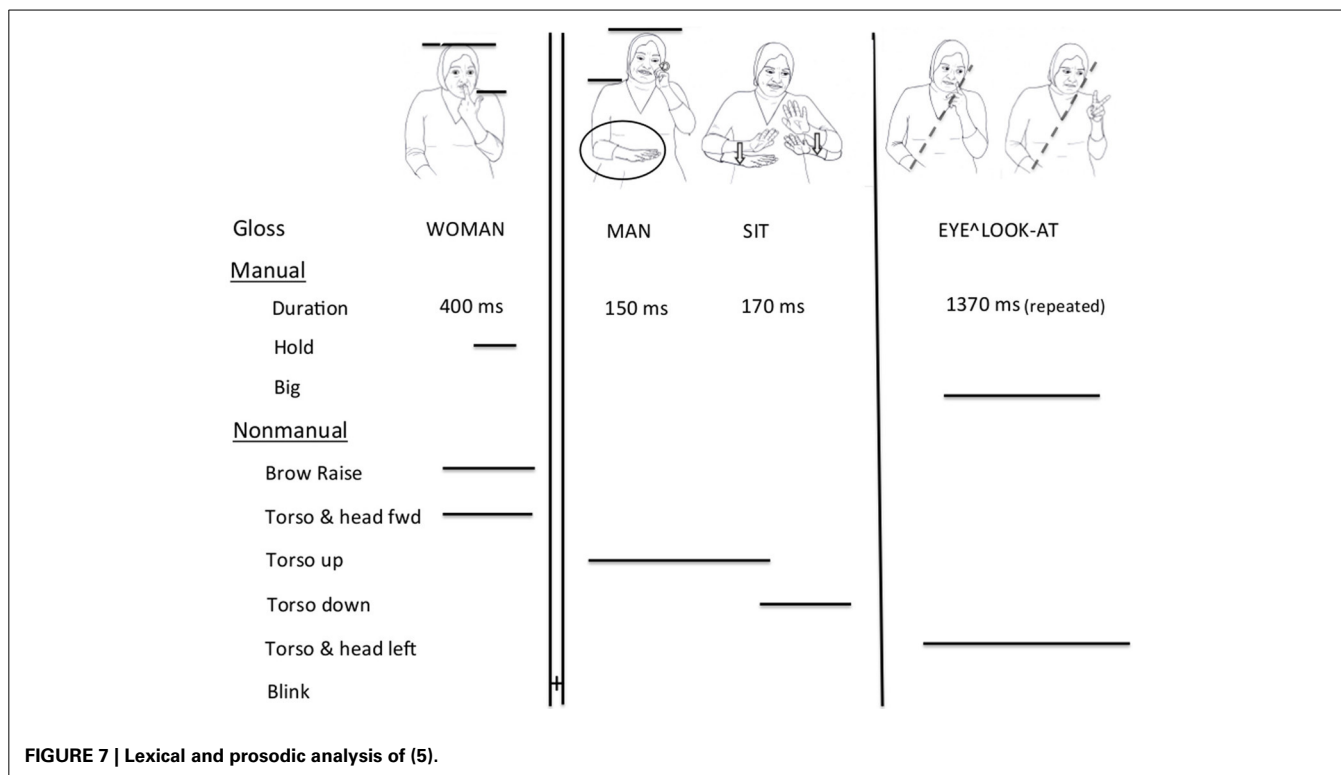
FIGURE 7 | Lexical and prosodic analysis of (5).

Third, LOOK-AT is reduplicated whereas SIT is not. Reduplication indicates aspectual marking in many sign languages (see Pfau et al., 2012 for a survey). The difference between LOOK-AT and SIT might be argued to be that of aspectual marking, that is, morphological rather than prosodic. However, in four of the 10 examples we provide below, the main predicate is not reduplicated, yet the difference in duration between the two predicates is noticeable.

We therefore attribute the difference in size and duration between the two predicates first and foremost to the difference in their prosodic positions, which are related to their constituency affiliation. Whether other factors, such as parts of speech and aspectual modulation, also play a role in differentiating between the two predicates, is an issue that we leave for future research.

Other instances of embedding

As noted above, after analyzing the response in (5), we noticed in our data that such reduced predicates occur in nine additional responses, and are used by three out of the six signers. In eight of these nine cases, as in the example analyzed above, the reduced predicate forms one prosodic constituent with a preceding noun, which denotes one of the human participants in the clips (a man, a woman or a girl); that is, there is no break in any of the prosodic signals between the predicate and the preceding sign, motivating our analysis that these signs belong to one prosodic unit[8]. In several cases, the facial expression and/or body posture that accompany the second sign, the modifier, spread regressively to the preceding noun, strengthening the impression that the two

signs form one prosodic unit. For example, in sentence (17), the signer raises her eyebrows for the sign GLASSES, but this facial expression starts on the preceding sign MAN. The result is that both signs are characterized by the same facial expression. The glosses of these 10 instances with reduced secondary predicates are provided in (9)–(18), where (12) contains two embedded predicates.

(9) WOMAN, [MAN SIT$_{170}$] EYE$^\wedge$LOOK-AT$_{400}$++. FACE LOOK++ (1.1.04)
'A woman looks at a sitting man. She looks at his face.'

(10) [MAN SIT$_{190}$] STAND-UP$_{320}$. STAND-UP. (1.1.12)
'A sitting man stands up. He stands up.'

(11) MAN, [ONE$^\wedge$WOMAN SIT$_{70}$], PICTURE SQUARE, LOOK SHOW$_{260}$. LOOK (1.1.14)
'A man shows a picture to a sitting woman. She looks.'

(12) [ONE$^\wedge$GIRL SMALL], [SIT$_{280}$ NERVOUS$_{250}$++ TEAR$_{350}$++ PAPER].TEAR$_{270}$ (1.1.15a)
'A seated nervous girl is tearing some paper', 'A girl, seated and nervous, is tearing paper.'

(13) [WOMAN HOUSE$_{110}$ THERE$_{90}$] RUN$_{590}$++ GO-PATH (1.1.18)
'A woman in the house is running.'

(14) [ONE$^\wedge$GIRL SMALL SIT$_{220}$], CRY$_{640}$++ ONE (1.1.30)
'A sitting girl is crying.'

(15) WOMAN1, WOMAN2 WOMAN1, [SIT$_{220}$ LEGS-CROSSED$_{350}$ ROLL$_{600}$++] (1.5.07)[9]

---

[8]Example (15) below is slightly different, as explained in fn. 9.

[9]In this example, the embedded predicate does not form one constituent with the preceding head noun. This might be a result of the fact that the head noun is a phrase consisting of three lexical nouns (WOMAN1, WOMAN2, WOMAN 1), and the signer is a bit hesitant as to which sign of "woman" to use. The embedded predicate also consists of two signs (SIT LEGS-CROSSED).

'A woman, sitting with her legs crossed, is rolling a ball.'

(16) MOTHER. IX MOTHER WOMAN SIT$_{420}$. GIRL BRUSH-OTHER$_{450}$++ BRUSH$_{300}$++, [MOTHER SIT$_{210}$] (1.5.22)
'A mother is sitting. A girl is brushing the sitting mother's hair with a brush.'

(17) MAN BOOK CUPBOARD [MAN GLASSES$_{270}$], BOOK PUT-IN$_{340}$ OPEN-CUPBOARD INTO (1.5.25)
'A man with glasses is putting a book in a cupboard.'

(18) [IX MAN GLASSES$_{160}$], BOOK PUT-IN$_{420}$. ARRANGE. PUT-IN (1.6.25b)
'A man with glasses is putting a book in. He's arranging things. He puts it in.'

We are now in a position to extend the analysis behind annotating (7) as (9) to examples (10)–(18). In all of these sentences, there are two signs that have a predicative function. But one of them is reduced in form and forms a constituent, either with a preceding nominal sign or, in the case of (12) and (15), with the following sign[10]. The phonological reduction is very clear. In the 10 cases, there was a marked difference between the duration of the two predicates: means of 214.8 ms for the short predicates, and 492.2 ms for the long predicates. Welch's two-tailed t-test suggests that this difference is robust, $t_{(14.89)} = 5.45$, $p < 0.001$. The phonological reduction, together with the prosodic cues indicating that the shorter predicate forms a constituent with the preceding nominal, are taken as evidence that the two predicative signs are of different grammatical status: the unreduced predicate is the main predicate in the clause, whereas the reduced predicate is a modifier of the preceding noun, an embedded predicate.

It is important to stress that not all responses containing two signs with a predicative function were analyzed as containing an embedded predicate. Compare clause (13) above (repeated here as 19) with (20);

(19) [WOMAN HOUSE$_{110}$ THERE] RUN$_{590}$++ GO-PATH (1.1.18)
'A woman in the house is running.'

(20) WOMAN, HOUSE$_{320}$, WALK-PATH, ONE (1.1.19)
'A woman is walking in a house.' or 'A woman, in a house, is walking'

The two responses are almost identical in terms of the signs used. Yet their prosodic structure is different, signaling different constituent structure and consequently a difference in the function of HOUSE in the two clauses. In (19), HOUSE is very short (110 ms), it is signed with a single movement (the two hands touch each other once), and there is no hold or pause between WOMAN and HOUSE. In (20), on the other hand, HOUSE is almost three times

longer, 320 ms. long. It has a double movement (the hands touch each other twice), and there is a hold on WOMAN, indicating a break between WOMAN and HOUSE. This break prevents us from interpreting HOUSE as forming a constituent with WOMAN; it forms its own prosodic constituent in the clause. Therefore, it is impossible to tell whether HOUSE modifies WOMAN or WALK-PATH, and both "the woman is walking in the house" and "the woman, in the house, is walking" are possible translations of this clause. Since there is no clear positive prosodic evidence for analyzing HOUSE as modifying WOMAN, we did not regard this response as an instance of embedding.

As these examples show, not all of the embedded predicates, that is, signs in a modifier position, are signs denoting actions or events. The signs that we found in this position are as follows: seven occurrences of stative predicates (SIT), one psych predicate (NERVOUS), one locative noun (HOUSE) and two instances of an attributive noun (GLASSES). They function as modifiers of the preceding noun, and were interpreted according to the nature of the modifier: "a sitting/seated man," "a nervous girl," "a woman in a house," "a man with glasses/a bespectacled man."

How should these predicates be analyzed? If we draw on the translations, it is tempting to analyze them as participles or reduced relative clauses: "a girl (who is) sitting," "a woman (who is) in the house," "a man (with/who has) glasses." Yet such a step is dangerous, since we are imposing the grammatical structure of one language (English) on another (KQSL). In English, participles have distinct morphological forms, and relative clauses are usually marked syntactically by a relative pronoun. In KQSL this is not the case. We have not found as yet any evidence for morphological markings that distinguish nouns from verbs, and no morphological evidence for the existence of participles. Furthermore (as is common in sign languages generally), we have not found any relativizers or other function words that mark subordination. It might be argued that the lack of evidence for such structures is due not to the simplicity in structure of the language but rather to the preliminary stage of investigation. However, studies of other village sign languages (e.g., ABSL, Aronoff et al., 2008; Padden et al., 2010, and other sign languages reported in Zeshan and de Vos, 2012) indicate that KQSL is not the exception; clear cases of syntactic manifestations of subordination have not been reported in other village sign languages. We are therefore reluctant to regard these reduced predicates as morphologically marked participles or bona fide subordinate clauses[11].

However, the prosodic cues clearly show that we have clausal constructions with two predicate signs, one of which, reduced in form, occurs in a modifier position, and is perceived as secondary, that is, as modifying a referent rather than as the main predicate.

Since both constituents consist of more than one sign, they each form a separate prosodic unit. However, the sign SIT is still very short, which we take to indicate that it is not the main predicate in the clause.

[10]In these two examples (12 and 15), the reduced predicate itself is being modified, by NERVOUS and LEGS-CROSSED respectively. In addition, the constituent referring to the subject in these examples contains more than one sign. The prosodic break after the subject constituent might be due to the relative "heaviness" of this constituent in these examples. However, we regard these clauses as instances of embedded predicates rather than two coordinate clauses (e.g. "a woman sits cross-legged and rolls a ball") because of the clear difference in duration and size between the predicates in the clauses.

[11]The literature on relative clauses in established sign languages suggests that in some languages relative clauses contain function words that occur at the boundaries of relative clauses. In some sign languages these signs are optional (e.g. the sign THAT in ASL, Liddell, 1978, and a sign referred to as "relative pronoun" in Turkish Sign Language, Kubus and Rathmann, 2011). In other sign languages the use of a function word as a marker of relative clauses is reported as "systematic," though the specific details and analyses vary across languages. See Pfau and Steinbach (2005) for German Sign Language, Branchini and Donati (2009) for Italian Sign Language, and Tang et. al. (2010) for Hong Kong Sign Language.

We therefore suggest that these are cases of *embedding* of a predicate within a clause, in the very basic sense of the term. In the next section we describe a possible path of emergence for this construction, and examine its consequences for our understanding of the development of embedding in language.

DISCUSSION: EMBEDDING BY HAND AND BY MOUTH

As mentioned earlier, subordination and embedding are prevalent in the languages of the world. From a functional point of view, subordinate and main clauses are construed as cognitively asymmetrical (Langacker, 1987; Cristofaro, 2003). Functional asymmetry is usually reflected in the asymmetry of the form. The dependent status of a subordinate clause is often manifested in the deranking of its predicate through the use of a predicate form that is not used in independent clauses. The predicate deranking can be realized in different ways, for example, by the use of nominal markers on dependent predicates (nominalization), or through the lack of formal distinctions characteristic of independent predicates (Stassen, 1985; Croft, 1991; Cristofaro, 2003).

Diachronic studies show that subordinate structures often originate from simpler structures. Deutscher (2000), for example, traces the emergence of subordinate clauses in Akkadian, documenting the transition from parataxis of two adjacent clauses to full embedding of one in the other. Yet the question of how this process originates has proven difficult to answer without data on novel embedding structures in a given language.

THE DIACHRONIC DEVELOPMENT OF SUBORDINATION

In their study of grammaticalization, Heine and Kuteva (2007) and Heine (2009) suggest that there are two main paths in which subordinate clauses arise[12]. The first is through a process of expansion, by reinterpreting a nominal as clausal. The second is by integration of two independent clauses into one, where one of the main clauses becomes subordinate to another. Heine (2009) notices that the first strategy usually gives rise to complement embedded clauses, while relative clauses and adverbial clauses usually arise via the second strategy, integration. Regarding the expansion process, Heine suggests that the first stage toward subordination is the appearance of a non-finite verb form (nominalization, infinitive or participle) in a nominal position in the clause. In subsequent stages, the phrases headed by a non-finite form of the verb acquire more and more verbal properties, eventually becoming clausal.

Deutscher (2009) points out that the scenario described by Heine and Kuteva (2007) and Heine (2009), though plausible, misses a crucial point. According to Deutscher, the real syntactic-cognitive feat of subordination is nominalization, the ability to derive a noun from a verb. The expansion of a nominalized verb into a clause is a secondary development, which builds on a structure that already contains subordination, at least from a cognitive point of view. As Deutscher puts it, "The ability of a language to derive a noun from a verb, that is, to reify a verbal predicate and to present it as a nominal argument or modifier, is at the core of

---

[12]The purpose of Heine (2009) is to show how devices that first served to structure independent sentences come to assume functions of subordination, and not necessarily to demonstrate the origin of subordination.

subordination." (p. 199). While Heine takes the transition from "Stage 0," which contains only nominal constituents, to "Stage 1," which contains a nominalized verb in a constituent position, as the first step of subordination, Deutscher argues that the transition from Stage 0 to Stage 1 is what needs to be explained, since it cannot be taken for granted. For him it is this step, the appearance of a nominalized verb in a non-predicative position, that needs to be explained. Yet most accounts of the development of subordination have neglected to do so.

Deutscher then points out that in the grammaticalization literature, it is very hard to find works on the development of markers that signal V > N change. He attributes the difficulty of finding such grammaticalization paths to the absence of source constructions with a nominal head that takes a verb as its complement. Such constructions necessarily involve nominalization, and therefore already contain an instance of subordination. He speculates that backformation might be one route to nominalization. He considers the French suffix -age which was originally attached to nouns: *mari* "husband" + *age* = "the state of being a husband." The denominal noun *mariage* was then reanalyzed as being derived from the verb *marier* "to marry," rendering -age a nominalizing suffix that can attach to verbs.

A different approach to the development of embedded structures is offered by Mithun (2009), examining data from Mohawk. Mithun suggests that in Mohawk integration of two clauses can be done only by prosodic means, with no syntactic indication of subordination. Two clauses, one a semantic complement of the predicate of the other, or two clauses that share an argument, may occur under one intonational contour, as in example (21) below (Mithun, 2009, p. 60):

(21) [*Tóka' ki' nè꞉'ne ki꞉ iakoia'takarénie's*]
    Maybe just it.is this bus
    [*t-hoti-ia'té-nha' wáhi'*],
    carried.them.here   TAG.QUESTION
      [*ki꞉ ratiksa'okòn꞉'a' thoné꞉non kèn꞉'en*]
      this children they.have.come here
    'Maybe the bus brought the children that came here'.

Mithun argues that the clauses in (21) are pronounced under a single prosodic contour. Within this overarching contour the pitch moves from High at the beginning of the first clause (*Tóka'*) to a full terminal fall at the end of the last clause (*kèn꞉'en*), with only partial pitch resets at the beginning of the internal prosodic phrases. She suggests that in this example one overall prosodic contour made up of small constituent sub-contours can be characterized as an instance of embedding. The author notices that the prosodic integration of the two clauses "reflects a kind of cognitive organization similar to that reflected by syntactic integration" (p. 61). That is, integration or embedding need not be reflected in the morpho-syntax; the cognitive feat of embedding is the ability to integrate two clauses in one construction. But this can be achieved by prosodic means alone in Mohawk. She further suggests that "The fact that we find prosodic structure without substantive syntactic structure suggests that prosodic structuring might, at least in some cases, precede syntactic structuring" (p. 61). Similar cases of subordination structures marked only

by intonation contours have been reported of other languages as well, e.g., Bambara (Bird, 1968, cited in Givón, 2012), and several languages in the Niger-Congo area (Givón, 2012).

These two approaches taken together suggest that embedding is first and foremost a cognitive operation, the integration of one proposition within another. This can be done by morpho-syntactic means, such as the development of nominal forms of verbs, or by prosodic means, by uttering two clauses under the same prosodic contour. KQSL differs from the languages investigated by Deutscher and Mithun in that it has neither morpho-syntactic marking of clause integration nor of parts of speech[13].

Studying the data from KQSL, we cannot claim that the secondary predicates we identified are nominal forms of verbs, since we have not discovered yet clear formal indication of parts-of-speech categorization in the language; all we can say is that they occur in a modifier position, function as modifiers, and have a form which can be regarded as less independent than the main predicates from phonological and prosodic points of view. However, we do think that KQSL offers a unique opportunity to look at the very first stage of embedding, the steps leading from Stage 0 to 1, that is, the possible initial stages of creating a subordinate predicate. Furthermore, since the only clues for the embedded status of these predicates is prosodic, KQSL bears witness to the role that prosody plays in the emergence of embedding.

A recent study on homesign (Hunsicker and Goldin-Meadow, 2012) may provide additional evidence as to how embedded constituents might arise in a communication system. The study describes the nominal constituents in the gestures of a young homesigner they call David. While David has iconic gestures ("nouns") and pointing gestures ("determiners"), he has also been developing more complex nominal constituents comprising of both a noun and a determiner. For example, the string in (22) could be parsed and interpreted in two ways: "that is a bird and it pedals," or the monoclausal "that bird pedals."

(22) point-at-bird BIRD PEDAL

What Hunsicker and Goldin-Meadow show is that the monoclausal interpretation is more plausible, implicating the existence of embedded structure: [[point-at-bird BIRD] PEDAL]. Here is

---

[13]The KQSL situation is somewhat reminiscent of initial stages of language development in children, when the parts-of speech-distinction is not fully established yet. Givón (2009), in a longitudinal study of the development of relative clauses, shows that children go through a stage in which they use various types of non-clausal post nominal modifiers, as in the following example:
(i) MOT: I like rabbits, don't you?
NIN: Yup. I like them. Like this one the red.
MOT: You like red rabbits?
NIN: Yup.
The adjective *red*, modifying the noun *rabbit*, fulfills a restrictive function and occupies the post-nominal slot that at a later stage is taken up by relative clauses. In Givon's opinion, those constructions may be considered early precursors of standard relative-clause forms. The KQSL examples are similar in that post nominal modifiers take the function of relative clauses without developing specific syntactic structures for marking relativization as yet. As a reviewer points out, English constructions such as *the man I saw* can also be considered cases of prosodically-marked embedding.

how the argument goes. First, they obtained the distribution of sentence lengths in David's production throughout the corpus without the sentences under examination, as measured in units (number of gestures). They then calculated two distributions of length: one with these sentences as "flat" structures, where each gesture corresponds to exactly one unit. In this case [[point-at-bird] [BIRD] [PEDAL]] would have three units. The other distribution was calculated with the hypothesized hierarchical structure, where [[point-at-bird BIRD] PEDAL] has two units. They found that the distribution with embedded structures provided a better fit to the data, leading them to conclude that the structures they investigate are to be seen as complex nominal constituents.

This case is remarkable since—as the authors discuss at length—David did not receive any structured hierarchical input from his caregivers. That he nevertheless developed embedded structure is in accordance with our findings as well; he managed to "squeeze" extra information into what was originally a basic nominal. One question that arises is how this was accomplished. We have provided evidence here that prosody is the resource used by KQSL signers to create embedded structure. Prosody is clearly co-opted in David's case as well: "Motoric criteria were also used to determine the end of a string of gestures and thus sentence boundaries. Two gestures were considered separate sentences if the child paused or relaxed his hands between the gestures. Gestures that were not separated by pause or relaxation of the hands were considered part of the same sentence." (Hunsicker and Goldin-Meadow, 2012, p. 736). The authors do not provide a prosodic analysis, and it is wise to be cautious about attaching labels such as "prosody" to generalizations about rate of signing in the homesign system of a young child. However, research on sign languages leads us to expect timing breaks in particular places. For example, consider example (23), adapted from (6b) in Hunsicker and Goldin-Meadow (2012, p. 743). If David's production is sensitive to factors such as signing rate and cognitive complexity of constituents, we would predict a short prosodic break between LONG and *point-downstairs*.

(23) [[point-at-self point-at-paddle LONG] point-downstairs]
     'my long paddle is there'

In sum, it might be possible that what regulates the length of sentences and constituents in David's system reflects the emergence of a prosodic system, which in turn enables hierarchical structures that may lead to embedding. Naturally, more work is required in order to substantiate this claim, in village sign languages as well as in homesign.

KQSL: A ROUTE TOWARD EMBEDDING

The structure in question—a clause containing a secondary predicate—appears to be a new phenomenon in KQSL, not just because of its limited frequency but also in light of the fact that 9 of the 10 instances documented were produced by the two younger signers in our pool, those aged 42 and 44. Labov (1972) famously discussed the notion of "apparent time": in the study of language variation and change, when two age groups vary on one sociolinguistic variable, it is likely that the older group

represents the previous stage in the development of the language, and the younger group, the later stage. It may well be the case that embedding is an emerging phenomenon in KQSL, owing to its relative prevalence among younger signers as compared to older ones. Under the assumption that the phenomenon described here is indeed a new development in the language, we speculate on where it might have come from and hypothesize about its possible implications for the future of KQSL.

We have argued above that the secondary predicates appear in a position designated for a nominal modifier, the modifier of a noun that forms a syntactic constituent with it. The recurrence of postnominal modifiers in this position creates a construction with a specific form and function, that is, a *slot* for postnominal modifiers. A crucial development in the language, then, is the emergence of the modifier slot. Though the presence of a modifier position may seem self-evident, findings from KQSL teach us that it is a significant development. In the data that we have so far, consisting of 213 responses for the video clips, and three conversations between dyads of signers (total of 25 min), there are very few clear instances of nominal modifiers. Though KQSL has words denoting properties, e.g., GOOD, BAD, FAST, NEW, they seem to function mostly as main predicates. The few clear examples of nominal modifiers are: PICTURE SQUARE ("a square picture"), GIRL FOUR ("four girls"), HOUSE SMALL ("small houses") and the forms FEMALE SMALL ("a girl") and MOTHER BIG ("a woman/ an adult woman"). Of these examples, only the last two are common in our data (about 70 occurrences), and we hypothesize that they are lexicalized collocations or compounds. These signs, then, clearly have a referential (and not a predicative) function, and they may have paved the way to the creation of a modifier slot[14].

Once a modifier slot is available, it can be used by various types of words. In our data we find locatives (WOMAN HOUSE THERE "a woman [who is] in a house there") and nouns denoting attributes (MAN GLASSES "the man with the glasses"). We also find signs denoting stative (stage-level) predicates such as NERVOUS and SIT.

We might speculate that a possible developmental path for the existing forms may have been moving from simple adjectives, such as "small" and "big," to locatives like "house" and then statives like "sitting." The next logical step would be the introduction of non-stative verbs such as perhaps "look" or "talk" in that slot. Such verbs can introduce complements, possibly leading to the appearance of clausal constituents in the modifier slot. However, this will require the development of more grammatical machinery. What signals predicates as secondary in our data is prosody: these signs tend to form one prosodic unit with the noun they modify, and are reduced in form. But if the embedded predicates take their own complements, they will probably form another prosodic unit, maybe separate from the head noun. In such a case, the language will need to develop a grammatical mechanism to mark this predicate and its complements as secondary. Some sign languages do this by means of facial expressions; in ISL, for example, relative clauses are often marked with a squint (Dachkovsky, 2008; Meir and Sandler, 2008; Dachkovsky and Sandler, 2009; Dachkovsky et al., 2013). Yet recruiting facial expressions for grammatical purposes takes time to develop; it is not there in early stages of a language (Sandler et al., 2011; Sandler, 2013). KQSL shows us that an initial step toward developing a relative clause is the creation of a modifier slot that can host different types of signs. The next step, the accommodation of clausal material such as arguments and adverbials in this slot, has not yet been attested.

CONCLUSION

This study has showcased a certain phenomenon in KQSL in order to shed light on the question of embedding in the evolution of natural language. Some younger adult signers of KQSL seem to use the postnominal modifier position as a slot open to several types of modifiers that can then act as secondary predicates. Locatives, concrete objects, psych predicates and stative predicates have been recruited as nominal modifiers. We have speculated that these developments might continue, with the embedded structures gradually allowing more elements and eventually leading to full-fledged relative clauses. Future work on KQSL will be needed to determine whether the hypothesis regarding the expansion of postnominal slots to host larger structures turns out to be correct.

Embedding, or, more specifically, recursion, has been at the core of recent debate on the nature of the human language faculty (see the debate in e.g., Hauser et al., 2002; Pinker and Jackendoff, 2005; Jackendoff, 2011; Watumull et al., 2014). While our work does not take a stand on this issue, it does provide evidence for an early stage of embedded structures in a natural language. Our data show that semantic embedding, that is, the embedding of a predicate within another, might be there from very early on in the development of a language, although the grammatical machinery to accommodate and mark embedding takes time to develop. Moreover, the earliest form of such grammatical machinery may not be clearly syntactic or morphological, but rather prosodic. The prosodic structure marks constituent boundaries of different degrees, signaling, inter alia, constituents within constituents, as in our case[15]. KQSL suggests that the initial stages of embedding may be quite modest: the creation of a slot for modifiers. Yet this modest step is crucial to get the wagon moving. Deutscher (2009, p. 212) argues that "Any attempt to explain the genesis of subordination can thus only begin to make sense if it explains the origins of nominalization, and if it shows how the ability to repackage a verb as a noun arises in contexts where it had not existed before." Our work shows how such repackaging might have started to develop. In a very young language such as KQSL, which hasn't developed morphological markings

---

[14]Deutscher (2005, appendix 3) speculates on the possible steps towards the emergence of a modifier slot in the development of a language. He suggests that deictic words such as demonstratives may have been the first words to have a modifier function, and by that they paved the way to create a modifier slot in the sentence structure. Our suggestion is compatible with Deutscher's hypothetical scenario, but we think that KQSL actually provides concrete data towards regarding the development of a modifier slot.

[15]As a reviewer points out, this classifies KQSL as a language with a "prosodic simple phrase grammar" in the terminology of Jackendoff and Wittenberg (in press).

for nominalization or syntactic mechanisms for subordination, such repackaging is done by means of prosody: predicates are "squeezed" into a modifier position, becoming embedded. The various predicates found in this structure further suggest that even such a humble step is composed of sub-steps, where more prototypical modifiers, such as *big/small*, may have paved the road to other, less typical, and even verbal predicates. KQSL enables us to zoom in on these initial stages, which are very hard to come upon when investigating the origins of subordination in spoken languages. The relative newness of sign languages compared to spoken languages makes them indispensable for our understanding of how linguistic structures arise and develop.

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://www.frontiersin.org/journal/10.3389/fpsyg.2014.00 525/abstract

REFERENCES

Al-Fityani, K., and Padden, C. A. (2010). "Sign language geography in the Arab world," in *Sign Languages: A Cambridge Survey*, ed D. Brentari (Cambridge: Cambridge University Press), 433–450.

Aronoff, M., Meir, I., Padden, C. A., and Sandler, W. (2008). The roots of linguistic organization in a new language. *Interact. Stud.* 9, 131–150. doi: 10.1075/is.9.1.10aro

Bird, C. S. (1968). Relative clauses in Bambara. *J. West Afr. Lang.* 5, 35–47.

Boyes Braem, P. (1999). Rhythmic temporal patterns in the signing of deaf early and late learners of swiss german sign language. *Lang. Speech* 42, 177–208. doi: 10.1177/00238309990420020301

Branchini, C., and Donati, C. (2009). "Relatively different: Italian Sign Language relative clauses in a typological perspective," in *Correlates Cross-linguistically*, ed A. K. Lipták (Hamburg: Benjamins), 157–191.

Crasborn, O., and Sloetjes, H. (2008). "Enhanced ELAN functionality for sign language corpora," in *Proceedings of LREC 2008, Sixth International Conference on Language Resources and Evaluation* (Marrakech).

Crasborn, O., van der Kooij, E., Waters, D., Woll, B., and Mesch, J. (2008). Frequency distribution and spreading behavior of different types of mouth actions in three sign languages. *Sign Language and Linguistics* 11, 45–67. doi: 10.1075/sll.11.1.04cra

Cristofaro, S. (2003). *Subordination.* Oxford, New York: Oxford University Press.

Croft, W. (1991). *Syntactic Categories and Grammatical Relations.* Chicago and London: The University of Chicago Press.

Dachkovsky, S. (2008). "Facial expression as intonation in Israeli Sign Language: the case of neutral and counterfactual conditionals," in *Signs of the Time, Selected Papers from TISLR 2004*, ed J. Quer. (Hamburg: Signum Verlag), 61–82.

Dachkovsky, S., Healy, C., and Sandler, W. (2013). Visual intonation in two sign languages. *Phonology* 30, 211–252. doi: 10.1017/S0952675713000122

Dachkovsky, S., and Sandler, W. (2009). Visual intonation in the prosody of a sign language. *Language and Speech* 52, 287–314. doi: 10.1177/0023830909103175

Deutscher, G. (2000). *Syntactic Change in Akkadian: The Evolution of Sentential Complementation.* Oxford: Oxford University Press.

Deutscher, G. (2005). *The Unfolding of Language.* New York, NY: Metropolitan Books.

Deutscher, G. (2009). "Nominalization and the origin of subordination," in *Syntactic Complexity: Diachrony, Acquisition, Neuro-Cognition, Evolution*, eds T. Givón and M. Shibatani (Amsterdam: Benjamins), 199–214.

Fenlon, J., Denmark, T., Campbell, R., and Woll, B. (2007). Seeing sentence boundaries. *Sign Lang. Linguist.* 10, 177–200. doi: 10.1075/sll.10.2.06fen

Givón, T. (2009). *The Genesis of Syntactic Complexity.* Amsterdam: Bemjamins.

Givón, T. (2012). "Toward a diachronic typology of relative clauses," in *Relative Clauses in Languages of Americas: A Typological Overview*, eds B. Comrie and Z. E. Fernandez (Amsterdam: Benjamins), 3–25.

Guerra Currie, A.-M. P., Meier, R., and Walters, K. (2002). "A cross-linguistic examination of the lexicons of four signed languages," in *Modality and Structure in Signed Language and Spoken Language*, eds R. Meier, K. Cormier, and D. Quinto-Pozos (Cambridge: Cambridge University Press), 224–236.

Hauser, M. D., Chomsky, N., and Fitch, W. T. (2002). The faculty of language: what is it, who has it, and how did it evolve? *Science* 298, 1569–1579. doi: 10.1126/science.298.5598.1569

Heine, B. (2009). "From nominal to clausal morphosyntax: complexity via expansion," in *Syntactic Complexity: Diachrony, Acquisition, Neuro-Cognition, Evolution*, eds T. Givón and M. Shibatani. (Amsterdam: Benjamins), 23–51.

Heine, B., and Kuteva, T. (2007). *The Genesis of Grammar: A Reconstruction.* New York, NY: Oxford University Press.

Hendriks, B. (2008). *Jordanian Sign Language: Aspects of Grammar From a Cross-Linguistic Perspective.* PhD dissertation, University of Amsterdam, Utrecht: LOT.

Herrmann, A. (2009). "Prosody in German sign language," *in Paper Presented at the Workshop on Prosody and Meaning*, Frankfurt am Main.

Hunger, B. (2006). Noun/verb pairs in Austrian Sign Language (ÖGS). *Sign Lang. Linguist.* 9, 71–94. doi: 10.1075/sll.9.1-2.06hun

Hunsicker, D., and Goldin-Meadow, S. (2012). Hierarchical structure in a self-created communication system: building nominal constituents in homesign. *Language* 88, 732–763. doi: 10.1353/lan.2012.0092

Jackendoff, R. (2011). What is the human language faculty? Two views. *Language* 87, 586–624. doi: 10.1353/lan.2011.0063

Jackendoff, R., and Wittenberg, E. (in press). "What you can say without syntax: a hierarchy of grammatical complexity," in *Measuring Grammatical Complexity*, eds F. Newmeyer and L. Preston (Oxford: Oxford University Press).

Jantunen, T. (2007). The equative sentence in finnish sign language. *Sign Lang. Linguist.* 10, 113–143. doi: 10.1075/sll.10.2.04jan

Johnston, T. (2001). Nouns and verbs in australian sign language: an open or shut case? *J. Deaf Stud. Deaf Educ.* 6, 235–257. doi: 10.1093/deafed/6.4.235

Johnston, T. (2003). "BSL, Auslan and NZSL: three signed languages or one?" in *Cross-linguistic Perspectives in Sign Language Research: Selected Papers from TISLR 2000*, eds A. Baker, B. van den Bogaerde, and O. Crasborn (Hamburg: Signum Verlag), 47–69.

Kimmelman, V. (2009). Parts of speech in Russian Sign Language: the role of iconicity and economy. *Sign Lang. Linguist.* 12, 161–186. doi: 10.1075/sll.12.2.03kim

Kisch, S. (2000). *Deaf Discourse: The Social Construction of Deafness in a Bedouin Community.* MA thesis, Tel-Aviv University [in Hebrew].

Kisch, S. (2007). Disablement, gender, and deafhood among the negev arab-bedouin. *Disabil. Stud. Q.* 27.

Kisch, S. (2012). "Demarcating generations of signers in the dynamic sociolinguistic landscape of a shared sign language: the case of language emergence: Al-Sayyid bedouin sign language," in *Sign Languages in Village Communities: Anthropological and Linguistic Insights*, eds U. Zeshan and C. de Vos (Berlin: Mouton De Gruyter and Nijmegen: Ishara Press), 127–152.

Kubus, O., and Rathmann, C. (2011). "Analysis of relative clause constructions (RCC) in Turkish Sign Language (TiD) in discourse modes," in *Talk Presented at 33. Jahrestagung der Deutschen Gesellschaft fur Sprachwissenschaft, Georg-August-Universität* (Göttingen).

Labov, W. (1972). "The social stratification of (r) in New York City department stores," in *Sociolinguistic Patterns*, ed W. Labov (Philadelphia, PA: University of Pennsylvania Press), 43–54.

Lanesman, S., and Meir, I. (2012a). "The survival of Algerian Jewish Sign Language alongside Israeli Sign Language in Israel," in *Sign Languages in Village Communities: Anthropological and Linguistic Insights*, eds U. Zeshan and C. de Vos (Berlin: Mouton De Gruyter and Nijmegen: Ishara Press), 153–179.

Lanesman, S., and Meir, I. (2012b). "Algerian jewish sign language—sociolinguistic sketch," in *Sign Languages in Village Communities: Anthropological and Linguistic Insights*, eds U. Zeshan and C. de Vos (Berlin: Mouton De Gruyter and Nijmegen: Ishara Press), 361–364.

Langacker, R. W. (1987). *Foundations of Cognitive Grammar:* Vol. I: *Theoretical Prerequisites.* Stanford, CA: Stanford University Press.

Liddell, S. K. (1978). "Non-manual signs and relative clauses in American Sign Language," in *Understanding Language Through Sign Language Research*, ed P. Siple (New York, NY: Academic Press), 59–90.

Liddell, S. K. (1980). *American Sign Language Syntax*. The Hague: Mouton de Gruyter.

Liddell, S. K., and Johnson, R. E. (1989). American sign language: the phonological base. *Sign Lang. Stud.* 64, 195–277. doi: 10.1353/sls.1989.0027

McKee, D., and Kennedy, G. (2000). "Lexical comparisons of signs from American, Australian, British and New Zealand Sign Languages," in *The Signs of Language Revisited: An Anthology to Honor Ursula Bellugi and Edward Klima*, eds K. Emmorey and H. Lane (Mahwah, NJ: Lawrence Erlbaum Associates), 49–76.

Meir, I. (2010). "The emergence of argument structure in two new sign languages," in *Syntax, Lexical Semantics and Event Structure*, eds M. Rappaport Hovav, E. Doron, and I. Sichel (Oxford: Oxford University Press), 101–113.

Meir, I., Israel, A., Sandler, W., Padden, C., and Aronoff, M. (2013). Community structure and language structure. *Linguist. Variation* 12, 247–291 doi: 10.1075/lv.12.2.04mei

Meir, I., and Sandler, W. (2008). *A Language in Space: The Story of Israeli Sign Language*. New York, NY: Lawrence Erlbaum Associates

Meir, I., Sandler, W., Padden, C., and Aronoff, M. (2010). "Emerging sign languages," in *Oxford Handbook of Deaf Studies, Language, and Education*, Vol. 2, eds M. Marschark and P. Spencer. (Oxford: Oxford University Press), 267–280.

Mithun, M. (2009). "Re(e)volving complexity: adding intonation," in *Syntactic Complexity: Diachrony, Acquisition, Neuro-Cognition, Evolution*, eds T. Givón and M. Shibatani (Amsterdam: Benjamins), 53–80. doi: 10.1075/tsl.85.03ree

Nespor, M., and Sandler, W. (1999). Prosody in Israeli sign language. *Lang. Speech* 42, 143–176. doi: 10.1177/00238309990420020201

Nespor, M., and Vogel, I. (1986). *Prosodic Phonology*. Dordrecht: Foris.

Nicodemus, B. (2009). *Prosodic Markers and Utterance Boundaries in American Sign Language Interpretation*. Washington, DC: Gallaudet University Press.

Nyst, V. (2012). "Shared sign languages," in *Sign Language. An International Handbook*, eds R. Pfau, M. Steinbach, and B. Woll (Berlin: Mouton de Gruyter), 552–574.

Ormel, E., and Crasborn, O. (2012). Prosodic correlates of sentences in signed languages: a literature review and suggestions for new types of studies. *Sign Lang. Stud.* 12, 279–315. doi: 10.1353/sls.2011.0019

Padden, C., Meir, I., Sandler, W., and Aronoff, M. (2010). "Against all expectations: Encoding subjects and objects in a new language," in *Hypothesis A/Hypothesis B: Linguistic Explorations in Honor of David M. Perlmutter*, eds D. Gerdts, J. Moore, and M. Polinsky (Cambridge, MA: MIT Press), 383–400.

Pfau, R., and Steinbach, M. (2005). Relative clauses in German sign language: extraposition and reconstruction. *Proc. North East Linguist. Soc.* 35, 507–521

Pfau, R., Steinbach, M., and Woll, B. (2012). "Tense, aspect, and modality," in *Sign Language. An International Handbook*, eds R. Pfau, M. Steinbach, and B. Woll (Berlin: Mouton de Gruyter), 186–204.

Pinker, S., and Jackendoff, R. (2005). The faculty of language: what's special about it? *Cognition* 95, 201–236. doi: 10.1016/j.cognition.2004.08.004

Sandler, W. (1999). "Cliticization and prosodic words in a sign language," in *Studies on the Phonological Word. Current Studies in Linguistic Theory*, eds T, Hall and U. Kleinhenz (Amsterdam: Benjamins), 223–255.

Sandler, W. (2011). Prosody and syntax in sign language. *Trans. Philos. Soc.* 108, 298–328. doi: 10.1111/j.1467-968X.2010.01242.x

Sandler, W. (2012). The phonological organization of sign languages. *Lang. Linguist. Compass* 6, 162–182. 10.1002/lnc3.326 doi: 10.1002/lnc3.326

Sandler, W. (2013). Dedicated gestures and the emergence of sign language. *Gesture* 12, 265–307. doi: 10.1075/gest.12.3.01san

Sandler, W., Aronoff, M., Padden, C., and Meir, I. (in press). "Language emergence: Al-Sayyid bedouin sign language," in *Cambridge Handbook of Linguistic Anthropology*, eds N. Enfield, P. Kockelman, and J. Sidnell (Cambridge: Cambridge University Press).

Sandler, W., and Lillo-Martin, D. (2006). *Sign Language and Linguistic Universals*. Cambridge: Cambridge University Press.

Sandler, W., Meir, I., Dachkovsky, S., Padden, C., and Aronoff, M. (2011). The emergence of complexity in prosody and syntax. *Lingua* 121, 2014–2033. doi: 10.1016/j.lingua.2011.05.007

Sandler, W., Meir, I., Padden, C., and Aronoff, M. (2005). The Emergence of grammar: systematic structure in a new language. *Proc. Natl. Acad. Sci. U.S.A.* 102, 2661–2665. doi: 10.1073/pnas.0405448102

Stassen, L. (1985). *Comparison and Universal Grammar*. Oxford: basil Blackwell.

Tang, Gladys, L. P., and Lee, J. (2010). "Strategies of relativatization in Hong Kong Sign Language," in *Paper Presented at the 10th International Conference on Theoretical Issues in Sign Language Research*, Purdue, IN.

Van der Kooij, E., Crasborn, O., and Emmerik, W. (2006). Explaining prosodic body leans in sign language of the netherlands: pragmatics required. *J. Pragmatics* 38, 1598–1614. doi: 10.1016/j.pragma.2005.07.006

Watumull, J., Hauser, M. D., Roberts, I. G., and Hornstein, N. (2014). On recursion. *Front. Psychol.* 4:1017. doi: 10.3389/fpsyg.2013.01017

Zeshan, U., and de Vos, C. (2012). *Sign Languages in Village Communities: Anthropological and Linguistic Insights*. The Hague: Mouton.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Lexical access in sign language: a computational model

Naomi K. Caselli* and Ariel M. Cohen-Goldberg

Department of Psychology, Tufts University, Medford, MA, USA

Psycholinguistic theories have predominantly been built upon data from spoken language, which leaves open the question: How many of the conclusions truly reflect language-general principles as opposed to modality-specific ones? We take a step toward answering this question in the domain of lexical access in recognition by asking whether a single cognitive architecture might explain diverse behavioral patterns in signed and spoken language. Chen and Mirman (2012) presented a computational model of word processing that unified opposite effects of neighborhood density in speech production, perception, and written word recognition. Neighborhood density effects in sign language also vary depending on whether the neighbors share the same handshape or location. We present a spreading activation architecture that borrows the principles proposed by Chen and Mirman (2012), and show that if this architecture is elaborated to incorporate relatively minor facts about either (1) the time course of sign perception or (2) the frequency of sub-lexical units in sign languages, it produces data that match the experimental findings from sign languages. This work serves as a proof of concept that a single cognitive architecture could underlie both sign and word recognition.

Keywords: neighborhood density, sign language, spreading activation, sub-lexical processing, sign perception, speech perception, lexical access

INTRODUCTION

One of the most important discoveries about language in the past half-century is arguably the fact that signed and spoken languages share fundamental aspects of their linguistic structure (Klima and Bellugi, 1979; Wilbur, 1979; Poizner et al., 1987; Lucas and Valli, 1992; Emmorey, 2002; Sandler and Lillo-Martin, 2006). The fact that all natural languages have common grammatical principles despite vast differences in modality has had critical implications for theories of the human language faculty and its evolution (e.g., Pinker, 1994; Jackendoff, 2002). Though a parallel line of research exists comparing the psycholinguistic mechanisms of signed and spoken language (Petitto et al., 2000; Sandler and Lillo-Martin, 2006; Emmorey et al., 2007; MacSweeney et al., 2008; Berent et al., 2013), much work remains. Far less is known, for example, about whether the mental lexicon is organized similarly across modalities and whether words and signs are activated and selected in similar ways. In the same way that the discovery of a common set of grammatical principles influenced theories of universal grammar, discovering similarities (or differences) in processing can profoundly advance our knowledge about psycholinguistic systems.

Within the psycholinguistic framework, the comprehension of a single word ultimately involves mapping a physical signal onto its meaning while the production of a single word involves the reverse process, mapping meaning to a physical signal. Multiple stages of processing have been posited to take place in between these two endpoints, most generally the identification (or in production, the preparation) of sub-lexical and lexical units (e.g., Dell, 1986; McClelland and Elman, 1986). According to a number of accounts, signed and spoken languages, should have similarly organized semantic systems (e.g., Jackendoff, 2012). At the same

time, their most peripheral elements clearly differ: signed languages utilize manual and facial articulators and are perceived through the visual system while spoken languages are produced with the oral articulators and are perceived through the auditory system.

There are a number of ways the language processing architecture could be organized with respect to these facts about the signed and spoken modalities. On the one hand, it's possible that signed and spoken languages utilize different cognitive mechanisms for all but the most central (i.e., semantic) stages of processing. It is also reasonable that a continuum of processing similarity could exist, where signed and spoken languages utilize similar cognitive mechanisms to achieve semantic processing but rely on increasingly different mechanisms to access the lexicon and process sub-lexical elements. Finally, it is also possible that identical psycholinguistic mechanisms underlie all stages of processing, with only the specific content differing across modalities (e.g., manual sign location vs. oral place of articulation).

In the present paper we consider the cognitive processes that underlie word and sign retrieval, that is, the mechanisms responsible for *lexical access*. We review the literature and find evidence that sign retrieval is influenced by factors that are specific to signed languages, suggesting that there may be modality-specific mechanisms for retrieving words and signs from the mental lexicon. Using a computational model, we explore the possibility that these differences are in fact superficial and that a common mechanism underlies lexical access in both modalities.

Computational modeling is a useful tool in the development of cognitive theories. In such an investigation, the modeler instantiates a particular cognitive theory in the code of a computer program. This encoding process is beneficial in and of itself

because it requires the modeler to state the theory in computationally explicit terms, defining its properties precisely. Once the theory has been translated thusly, the modeler may then use the program to test the theory. By running the program, the modeler runs a simulation of the theory, obtaining specific outputs for specific inputs. This allows the modeler to determine the predictions of the theory (e.g., in lexical access, if a sign's basic components are activated in this sequence, what are the consequences for the sign's activation?). This can be especially important in complex systems where it may be otherwise difficult to determine how the system will function (e.g., how are signs activated in a system with many connections and feedback loops?). Finally, the modeler compares the predictions generated by the simulation to empirical data. To the extent that the behavior of the simulation matches human behavior, we can conclude that the principles that underlie human behavior might be the same as those that underlie the model (see McCloskey, 1991, for a discussion of the difficulties in assigning credit and blame in simulations). Failure to capture empirical performance, by contrast, would provide an argument that the theory instantiated by the computer program is not an accurate description of human cognition (e.g., Goldberg and Rapp, 2008). Like laboratory experiments, most simulation work focuses on explicating a particular aspect of a cognitive domain. In this pursuit, simulations typically systematically vary the property of interest while keeping extraneous factors constant, either by using constant values or by not modeling the property at all. The advantage of this approach in modeling and in laboratory experiments is that it is possible to isolate the effects of variables of interest, though it reduces the ecological validity of the study. Nevertheless, simulation can form an important role in the feedback loop of theory building (Peschl and Scheutz, 2001).

In the present paper, we develop a computational simulation of sign access that imports core access principles that were developed specifically to account for phenomena observed in spoken (and written) lexical access (Chen and Mirman, 2012). The strength of this model in the present case is that it contains no elements that are specific to signed or spoken languages, allowing us to determine if an abstract set of principles is capable of accounting for lexical access across modalities. We show that if a model containing these core principles is elaborated to incorporate relatively minor facts about either (1) the time course of sign perception or (2) the frequency of sub-lexical units in sign languages, it produces data that qualitatively match the experimental findings from sign languages. We argue that these simulations serve as an existence proof, demonstrating that a single computational mechanism could in theory be responsible for lexical access in signed and spoken languages. Finally, we use the simulation to generate a novel prediction about how lexical access is accomplished in sign language that we hope spurs future research.

In spoken word processing, one of the most well-documented findings is that the degree to which a word is phonologically related to other words influences how that word is processed. In spoken and written language, neighborhood density, a measure of how interconnected a given word is, has been typically been defined as the number of words that differ from the target word by one grapheme or phoneme (Coltheart et al., 1977;

Luce and Pisoni, 1998). Psycholinguistic research has demonstrated that neighborhood density influences speech perception, speech production, and written word perception, but the effect differs by task and modality. In spoken production neighborhood density is facilitatory (Vitevitch, 1997, 2002; Mirman et al., 2010 though recent studies have suggested a more complicated picture: Mirman and Graziano, 2013; Sadat et al., 2014) while in spoken perception neighborhood density is inhibitory (e.g., Goldinger et al., 1989; Dufour and Peereman, 2003). In visual word recognition neighborhood density is facilitatory (Andrews, 1992), except for high frequency words in which case neighborhood density is inhibitory (e.g., Grainger et al., 1989; Davis et al., 2009)[1].

Until recently, the theoretical accounts of these neighborhood density effects have differed depending on the modality. For example, in speech perception neighbors were posited to be inhibitory because multiple candidate words compete for selection (McClelland and Elman, 1986), while in speech production neighbors were thought to be facilitatory because of the dominant influence of feedback connections (Dell and Gordon, 2003). Chen and Mirman (2012) proposed a single architecture that attempts to unify the pattern of reversals in *spoken and written language*. At the heart of their architecture is a spreading activation system with two kinds of connections between linguistic units: inhibitory lateral connections between lexical items and facilitatory "vertical" connections between lexical items and phonemes/graphemes and between lexical items and semantic units (see **Figure 1A**). Vertical connections are bidirectional, allowing for the feedforward as well as feedback flow of activation, while lateral connections are unidirectional, meaning that two lexical items can inhibit each other with different strengths. The system differs from a standard spreading activation architecture in that the strength of a lexical unit's inhibitory connections to other units varies as a function of the unit's activation. Rather than being fixed, inhibitory weights vary according to a sigmoid function: if the unit's activation is low the weight on the inhibitory connection is small; if the unit's activation is high the weight is large (see **Figure 1B**).

Lexical items thus send both facilitatory *and* inhibitory activation to other lexical items. For example, imagine an individual hears the word *cat*. As phonetic information is translated to phonological information, the matching sub-lexical units /k/, /æ/, and /t/ become active. As sub-lexical units receive activation, they each send activation through feedforward connections to the target word and its neighbors (*cap*, *sat*, *cot*, etc.). As the lexical items become active, they feed activation back to the sub-lexical units, which in turn feed activation forward, facilitating the target and its neighbors. At the same time, as the target and neighbors become active they inhibit each other through lateral

---

[1]A related reversal has been shown for semantic neighbors (words that are semantically but not phonologically related to the target). Neighbors that share many semantic features with the target inhibit processing while neighbors that share few features facilitate target processing (Mirman and Magnuson, 2008). As the simulations presented here model form ("phonological") neighbors in sign language processing, we focus the remainder of the review on the literature in spoken word and sign language processing rather than reading or semantics.
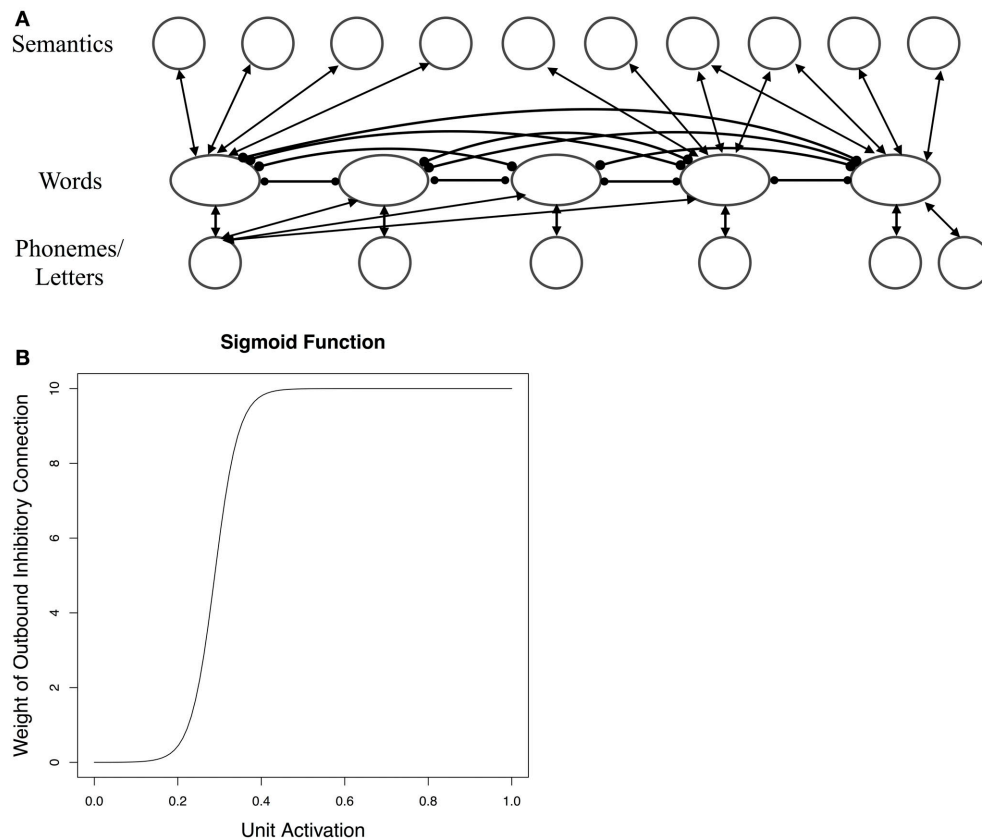
FIGURE 1 | Chen and Mirman (2012) Architecture. Panel **(A)** illustrates the spreading activation architecture used by Chen and Mirman (2012) to account for the pattern of reversals of neighborhood density effects in spoken and written language. Facilitatory connections are drawn with arrows, and inhibitory connections are drawn with circle endpoints. In this architecture, as demonstrated in panel **(B)**, the amount of inhibition a given lexical item exerts is scaled by a sigmoid function of its activation. Figures adapted from Chen and Mirman (2012).

(lexical-lexical) connections. Neighbors thus simultaneously activate and inhibit the target word.

Chen and Mirman suggest that the reversals in the direction of neighborhood density effects observed in spoken and written language result not from architectural differences across modalities but from delicate shifts in the balance between the facilitation and inhibition sent by a word's neighbors. When a neighbor is strongly activated, the amount of inhibition it sends outweighs the amount of facilitation it sends, due to the activation-dependent weighting of the inhibitory connections (high activation results in a large inhibitory weight). The net effect on the target item is inhibition. Conversely, when a lexical item is weakly activated, the amount of facilitation it sends outweighs the inhibition, resulting in facilitation of the target word. To generalize, strong neighbors inhibit while weak neighbors facilitate. According to their argument, differences in the task being performed lead to shifts in net facilitation or inhibition, causing neighbors to inhibit spoken recognition but facilitate spoken production. Specifically, neighbors become highly activated during speech perception (and thus have an inhibitory influence) since they are directly activated by sub-lexical units (/k/ /æ/ activate both *cat* and *cap*). By contrast, neighbors are relatively weak in production since the only activation they receive is through

feedback from sub-lexical units (*cat* sends feedback activation to /k/ and /æ/, which in turn activate *cap*).

Turning to signed language, sign processing in many ways is like word processing. Like words, signs are accessed automatically (Dupuis and Berent, 2013). Phonological structure is one of the core organizing properties of all languages, including sign languages (Goldin-Meadow et al., 1995). Like the sounds in words, signs are composed of discrete meaningless formal units such as hand configuration or location[2]. As in spoken language, lexical access in signed language is thought to entail a two-step procedure involving sub-lexical and lexical levels of processing in production (Thompson et al., 2005; Corina and Knapp, 2006a; Baus et al., 2008) and perception (Corina and Emmorey, 1993; Corina and Hildebrandt, 2002; Mayberry and Witcher, 2005; Dye and Shih, 2006; Carreiras et al., 2008; Carreiras, 2010).

Far fewer studies have examined the role of "phonological" (formal) neighbors in sign language, though the emerging pattern

---

[2]Early literature proposed 4 classes of sub-lexical units or "parameters": hand-shape, location, movement, and palm orientation (Stokoe, 1972). Recently, more nuanced systems have been proposed for describing signs (Sandler, 1989; van der Hulst, 1993; Brentari, 1998; van der Kooij, 2002) though Stokoe's four parameters remain prevalent in the psycholinguistic literature.

is that neighbors also influence sign processing. To date, neighbors in sign language have generally been defined differently than they have been defined in spoken language. Rather than defining neighbors as signs that *differ* by one sub-lexical unit (minimal pair neighbors), neighbors have been defined as signs that *share* one sub-lexical unit (though other definitions have also been used: Mayberry and Witcher, 2005; Corina and Knapp, 2006a; Dye and Shih, 2006). Signs that share the same handshape are typically referred to as "handshape neighbors," signs that share the same location are called "location neighbors," and so on. Though this approach makes comparison between signed and spoken language somewhat difficult, it has been used in part because there are far fewer minimal pairs in sign languages relative to spoken languages (van der Kooij, 2002).

This approach has revealed that the effect of neighborhood density in sign perception differs depending on the *specific type* of neighbor. In a study of Spanish Sign Language (LSE) processing, Carreiras et al. (2008) found that signs with many *handshape* neighbors (having "dense handshape neighborhoods") are easier to identify in a lexical decision task than signs with few handshape neighbors. Meanwhile, signs with dense *location* neighborhoods are harder to identify than signs with few location neighbors. Inhibitory effects have also been observed in primed lexical decision tasks in American Sign Language (ASL), where location primes inhibit target processing (Corina and Emmorey, 1993; Corina and Hildebrandt, 2002)[3]. Finally, a similar pattern has been observed in production. In a picture-sign interference task, Catalan Sign Language (LSC) signers named pictures more slowly when the to-be-named picture was presented alongside a distracter sign that used the same location and more quickly when the distracter shared the same handshape or movement (Baus et al., 2008).

It is important to note that these effects have not been universally found. Some studies have failed to find priming effects with either handshape neighbors (Corina and Emmorey, 1993; Dye and Shih, 2006) or location neighbors (Dye and Shih, 2006)[4] though there is some suggestion that these null effects may be due to varying ISI and insufficient power (see Carreiras, 2010). Similar null effects of location neighbors and handshape neighbors have been documented in production as well (Corina and Knapp, 2006a). There is also some evidence that the effects of neighbors may be modulated by language experience. In the only known study to define neighbors in the same way as spoken language, Mayberry and Witcher (2005) found facilitatory neighborhood effects for signers who started learning ASL between ages 4 and 8, inhibitory effects for signers who started learning ASL between the ages of 9 and 13, and no effects for signers who learned ASL from birth. Clearly more research is needed but to summarize, when neighbors have been defined as signs that share one feature with the target, the studies that have found

significant effects have consistently indicated that location neighbors inhibit lexical access while handshape neighbors facilitate access.

Putting these findings together, we see that in spoken language it is the specific task (perception vs. production), while in signed language it is the specific type of neighbor (location vs. handshape) that determines facilitation and inhibition. How might we account for these differences? One possibility is to assume that there are different computational principles at work in signed and spoken language, leading to fundamental differences in the way words and signs are activated during language processing (e.g., Corina and Knapp, 2006b; Baus et al., 2008). The fact that it matters in sign language whether a neighbor shares its location or its handshape with the target suggests that there are sign language-specific retrieval mechanisms since there is no exact corollary of these parameters in spoken language. These different mechanisms could have their origins in the different neural substrates that may underlie signed and spoken word processing. For example, the difference between location and handshape in sign processing may be due to the fact that spatial location and object recognition are carried out via different neural "streams" in the visual system (e.g., Mishkin et al., 1983). The different mechanisms could also arise because handshapes are compositionally more complex than locations since they comprise many features (selected fingers, abduction, etc.) while locations can be specified by a single feature (e.g., *shoulder*; Corina and Knapp, 2006b). Another difference is that handshape is perceived categorically, while location is not (Emmorey et al., 2010). These sorts of explanations imply that the language architecture differs across the modalities.

Another possibility is that spoken and signed languages make use of the same core mechanisms to access the mental lexicon and it is a handful of relatively peripheral differences between modalities that accounts for the differences in the way neighbors affect processing. Chen and Mirman's theory of lexical access accounts for the pattern of reversals observed in spoken (and written) language with a single core lexical access mechanism, varying only the most peripheral elements across modality (the sequence of activation of sub-lexical units in speech perception and word recognition). In the same way, it could be the case that the same computational mechanism underlies sign and word processing and the pattern of reversals apparent in sign language is a result of variation in the peripheral facts about location and handshape in signs. To the point, location neighbors may be inhibitory and handshape neighbors facilitatory because facts about sign locations and handshapes may make location neighbors stronger competitors than handshape neighbors.

In the present investigation, we explore three reasons that location neighbors might generally be stronger competitors than handshape neighbors. The first possibility relates to the temporal order of a sign's perception. As a sign unfolds over time, location is identified ∼30 ms earlier in perception than handshape (Grosjean, 1981; Emmorey and Corina, 1990, though see Morford and Carlson, 2011). This might mean that location sublexical units send activation to neighbors for a relatively long time, enabling location neighbors to become strong competitors. By the same token, the later recognition of handshape might mean that

---

[3]Corina and Hildebrandt (2002) found marginally significant inhibitory effects of location primes.

[4]Note that Dye and Shih (2006) found a facilitatory effect of primes that shared both movement and location. However, because targets and primes shared two sub-lexical units, it is difficult to know whether the source of the effect was location, movement, or an interaction of the two.

handshape sub-lexical units become activated later in time and send activation to neighbors for only a relatively short amount of time, leading handshape neighbors to become only weakly activated. It is thus possible that the timing of sub-lexical feature activation in perception is what causes location neighbors to be inhibitory and handshape neighbors to be facilitatory in recognition.

The second possibility relates to the absolute number of neighbors a target sign has. Although Carreiras et al.'s (2008) design crossed neighbor type (location/handshape) with density (high/low), the number of neighbors in the high and low density conditions varied across neighbor type. Specifically, the high density location neighborhoods were almost seven times larger on average than the high density handshape neighborhoods. It could be simply that the purported difference between location and handshape neighborhoods was actually due to the difference in neighborhood size across the location and handshape conditions. That is, it is possible that a large number of neighbors (e.g., the number of neighbors in the location condition) inhibits perception, but a "medium" amount of neighbors (e.g., the number of neighbors in the handshape condition) facilitates perception. According to this hypothesis, it is the absolute number of neighbors that causes location neighbors to be inhibitory and handshape neighbors to be facilitatory in recognition.

The last possibility is that location is more robustly represented than handshape. There is a wealth of evidence that this may be the case. Location is misperceived less frequently than other features (Orfanidou et al., 2009), and is easier to remember than movement and orientation (Thompson et al., 2005). Location errors are less frequent than handshape errors (Klima and Bellugi, 1979; Corina, 2000; Hohenberger et al., 2002), and location is learned sooner (e.g., Marentette and Mayberry, 2000). If location representations are more robust than handshape representations, location *neighbors* will become strongly activated during sign recognition while handshape neighbors will be relatively weakly activated. Within the Chen and Mirman architecture, this would cause location neighbors to have a net inhibitory effect and handshape neighbors to have a net facilitatory effect on target recognition.

There are several reasons that location may be more robustly encoded than handshape, for example, locations might be more salient, draw more attention, or be attended to at an earlier age than other sign parameters. For the purposes of this investigation, we focus on a possibility that arises because of the particular way that neighbors have been defined in sign language research. When neighbors are defined as signs that share *one* sub-lexical unit rather than signs that share all but one sub-lexical unit (as in spoken and written language research), neighborhood density is actually the same as *sub-lexical frequency*. What Carreiras et al. (2008) called an effect of neighborhood density—a lexical property—could actually be an effect of sub-lexical frequency. In their stimuli, the average location was seven times more frequent in the language than the average handshape. We consider the possibility that sub-lexical frequency (or other factors, such as salience/attention) influences how robustly sub-lexical units are encoded, which we instantiate as different

levels of resting activation. According to this proposal, high frequency sub-lexical units (locations) could have high resting levels of activation leading location neighbors to become strong (inhibitory) competitors. Low frequency sub-lexical units (handshapes) could have low resting levels of activation, leading handshape neighbors to become weak competitors and result in net facilitation.

We report the results of 3 simulations of sign recognition using a lexical network that utilizes the activation principles proposed by Chen and Mirman (2012) and that incorporates differences in sub-lexical activation and timing and neighborhood density, as described above. The use of computer simulations allows us to test how sign perception could function in a system that has no intrinsic location or handshape, or any other sign-specific features. We can test whether the factors that influence the strength of a neighbor's activation described above are sufficient for obtaining the observed pattern of facilitation and inhibition. If the simulations are capable of reproducing the observed effects, they will serve as a proof of concept that language-general principles are sufficient to account for lexical access in sign language. If the simulation is incapable of reproducing the empirical results, we conclude that sign access involves different—i.e., sign language-specific—retrieval mechanisms than spoken language (though null results are always difficult to interpret).

MODEL ARCHITECTURE

Like Chen and Mirman (2012), the structure of the architecture comprised two layers of units: a sub-lexical level and lexical level (see **Figure 2**). Bidirectional facilitatory weights connected the lexical and phonological levels, and unidirectional lateral inhibitory weights connected lexical items (see **Table 1** for parameter values). As in Chen and Mirman (2012) lateral inhibitory connections were scaled by a sigmoid function of word activation that forces rapid selection of only one lexical item (in all models $\beta = 35$ and $x_0 = 0.3$, following Chen and Mirman):

$$y = \frac{15}{1.5 + e^{-\beta(x-x_0)}}$$

In order to simulate the recognition of a single target sign, the sub-lexical units associated with the target were activated through external input, and the activation of the target sign was taken as a measure of lexical access. The simulations reported here orthogonally varied the timing (Simulation 1) and amount of activation given to the sub-lexical units (Simulation 2) as well as the number of neighbors shared by the target (Simulation 3). We provide the details of these manipulations in the simulations below. Note that we modeled average reaction times for each cell (density: high and low; neighbor type: handshape and location) rather than reaction times for particular items. The assumptions regarding timing, sub-lexical frequency, and neighborhood density were also derived from averages rather than particular lexical items. The net effect of a neighbor on the target was calculated by subtracting the activation of a target no neighbors from the activation of the target with a neighbor (or neighbors). The simulations presented here were implemented using PDPtool in MATLAB (McClelland, 2009).
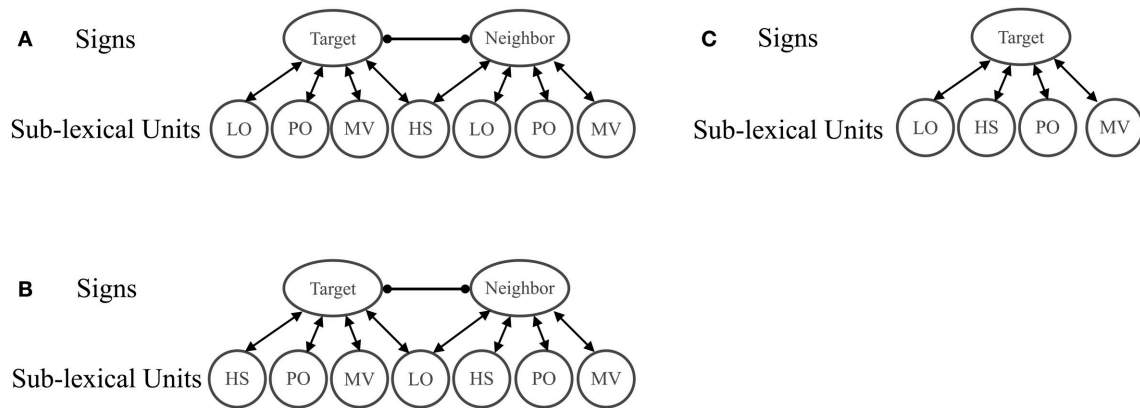
FIGURE 2 | Model Architecture. Activation of the target with a handshape neighbor **(A)** or location neighbor **(B)** was compared to activation of the target without a neighbor **(C)**. Neighbors were considered to have a facilitatory effect on sign recognition if the target item with a neighbor (**A,B**) became active more quickly than the target item without a neighbor **(C)**. Neighbors were considered to have an inhibitory effect if the target item with a neighbor became active more slowly than the target item without a neighbor.

Table 1 | Values Used in All Simulations.

Parameter	Value
Sub-lexical unit to sign excitation	0.2
Sign to sub-lexical unit excitation	0.2
Sign to sign inhibition	See formula
Resting activation	0 unless otherwise specified
Sub-lexical unit Decay	0
Word Decay	0

SIMULATION 1: TIMING

In Simulation 1, we tested the hypothesis that the effects of location and handshape can actually be attributed to the sequence with which sub-lexical units become active in perception. To do this, we manipulated the timing of the activation of the sub-lexical units in accordance with the average time of sub-lexical unit identification from behavioral data. Emmorey and Corina (1990) report that location and orientation are identified first (146 ms on average), followed by handshape (172 ms), and then movement (238 ms). To simulate timing, two of the target sub-lexical units ("location" and "orientation") received input for 3 cycles (equivalent to ~30 ms) before the "handshape" sub-lexical unit was activated for 7 cycles (equivalent to ~70 ms). Finally, the "movement" sub-lexical unit was activated for the remaining cycles. The effect of having a location neighbor was simulated by creating an additional lexical unit that shared the location unit with the target but had distinct orientation, handshape, and movement features (see **Figure 2A**). The effect of having a handshape neighbor was simulated the same way, except that the neighbor shared the handshape unit with the target (see **Figure 2B**). Since we are simulating the recognition of the target item, only the target's sub-lexical units received activation—none of the neighbor's sub-lexical units were activated except for the shared unit. The amount of external input applied to the sub-lexical units was set to 2, though we explored other levels of activation and the results were qualitatively the same throughout.

SIMULATION 1 RESULTS

The results of Simulation 1 are presented in **Figure 3**. As predicted, when the shared sub-lexical unit became active early in processing (as is empirically the case with location), the neighbor contributed net inhibition to the target sign. When it became active late in processing (as has been demonstrated for handshape), the neighbor contributed net facilitation to the target sign. The fact that the network tested in Simulation 1 produced the correct pattern of behavior suggests that the inhibition and facilitation observed for location and handshape neighbors in sign recognition may be due to differences in when different sub-lexical units are activated in perception.

SIMULATION 2: SUB-LEXICAL FREQUENCY

In Simulation 2, we tested the hypothesis that the effects of location and handshape could actually be due to differences in how robustly encoded the sub-lexical units are. We simulated this possibility by manipulating the resting level of activation of the sub-lexical units in accordance with the average sub-lexical frequencies of the location and handshape parameters. As described above, in the existing behavioral research the high density location neighborhoods ($M = 203$, range = 203–203) were almost seven times larger than the high density handshape neighborhoods ($M = 28$, range = 21–35; Carreiras et al., 2008). To model this difference, the resting activation of one sub-lexical unit (the "location" unit) was set to 0.7 while the resting level of the other units was set to 0.1. The amount of external activation applied as input to the sub-lexical units was set to 1, though the results are qualitatively the same with other levels of input. All sub-lexical units received external activation simultaneously, rather than sequentially as in Simulation 1. We note that resting level of activation is only one way of modeling frequency (Dahan et al., 2001; Knobel et al., 2008), and resting activation could also be thought to correspond to attention or salience (e.g., Mirman et al., 2008).
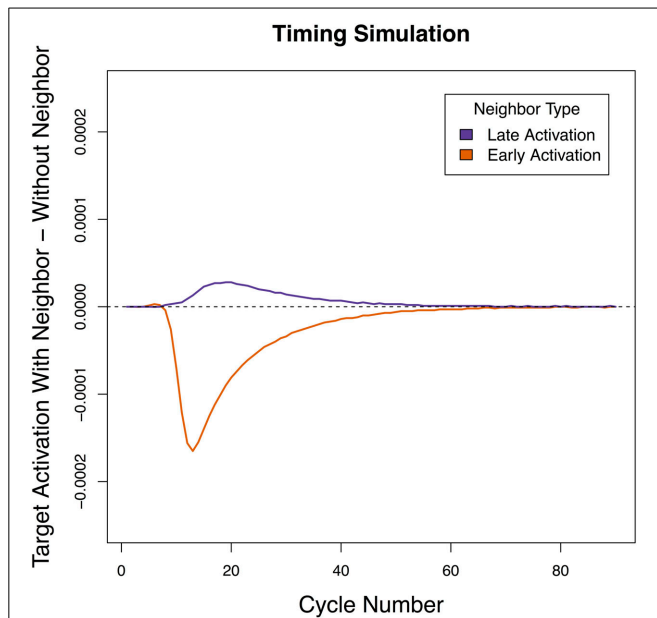
FIGURE 3 | Net contribution of a handshape neighbor and a location neighbor when the timing of sub-lexical sub-lexical unit activation was manipulated. Handshape neighbors had a net facilitatory effect on the target, while location neighbors had a net inhibitory effect on the target.
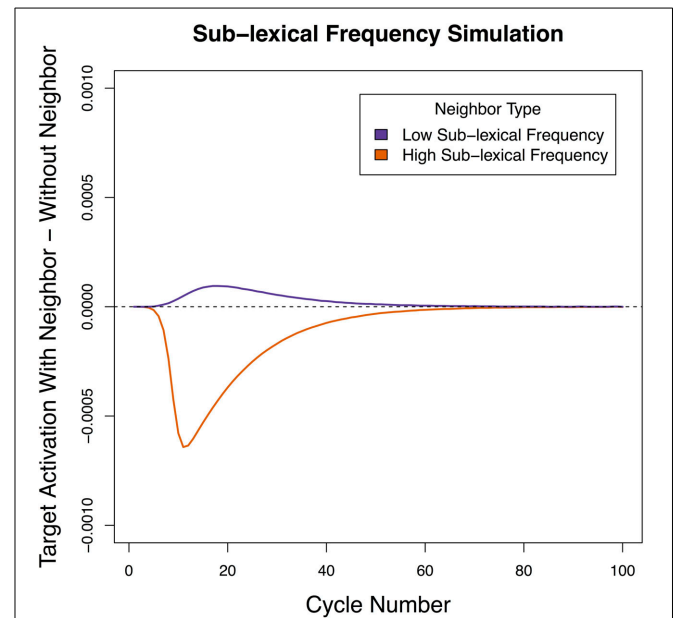

FIGURE 4 | Net contribution of a handshape neighbor and a location neighbor when the resting level of activation of sub-lexical units was manipulated. Handshape neighbors had a net facilitatory effect on the target, while location neighbors had a net inhibitory effect on the target.

SIMULATION 2 RESULTS

As in Simulation 1, Simulation 2 revealed that a when the shared feature had high resting activation the neighbor contributed net inhibition to the target sign, and when the shared feature had low resting activation (which corresponded to handshape) the neighbor contributed net facilitation to the target sign (see **Figure 4**). The results were qualitatively the same within ±0.2 units of resting activation. This suggests that facts about sub-lexical frequency could be responsible for the patterns of facilitation and inhibition in sign recognition.

INTERIM DISCUSSION

Both simulations demonstrated that it is possible to model the pattern of reversals seen in behavioral studies of sign perception with minimal modifications to the architecture thought to underlie spoken language. At the sub-lexical level, varying either the timing of activation or the amount of resting activation is sufficient to produce quantitatively similar patterns to what has been observed with humans performing sign recognition. These results demonstrate that differences in the timing with which location and handshape targets are perceived and differences in the robustness with which these parameters are encoded (as modeled using sub-lexical frequency) are computationally tractable explanations for the pattern of reversals in sign language.

SIMULATION 3: NUMBER OF NEIGHBORS

The first two simulations evaluated whether manipulations of sub-lexical properties can produce the observed pattern of facilitation and inhibition. In Simulation 3 we consider whether the pattern of reversals is due to activity at the lexical level, in particular the number of neighbors that are active during processing.

Two conditions were simulated: having a high neighborhood density (HND) and having a low neighborhood density. In the HND condition, which simulated the size of the location neighborhoods in Carreiras et al. (2008), there were four neighbors and in the low neighborhood density condition (LND; simulating the handshape neighborhoods), there was only one neighbor (see **Figure 5**). To determine the net contribution of the neighbor(s), the activation of the target in the LND condition (**Figure 5B**) and the HND condition (**Figure 5A**) was compared to the activation of the target without a neighbor (**Figure 5C**). To test the generality of the density effects, we tested LND and HND conditions using different amounts of external activation to the target sublexical units. We report data for external activation levels of 1 and 9 but the results are qualitatively the same at other input levels. In order to isolate the effect of lexical neighborhood density, all sublexical units simultaneously received the same amount of external activation.

SIMULATION 3 RESULTS

A very different pattern emerged in Simulation 3 than the previous 2 simulations. Here, neighborhood density did not determine the direction of the effect (the HND and LND conditions patterned together) and what determined whether the effect was facilitatory or inhibitory was the amount of activation applied to the input units (**Figure 6**). Specifically, when a low amount of activation was applied, both HND and LND were facilitatory and when a high amount of activation was applied, both HND and LND inhibitory. In all cases, having four neighbors magnified the effect of having a single neighbor—when a single neighbor was facilitatory, four neighbors were more facilitatory, and when a single neighbor was inhibitory, four neighbors were
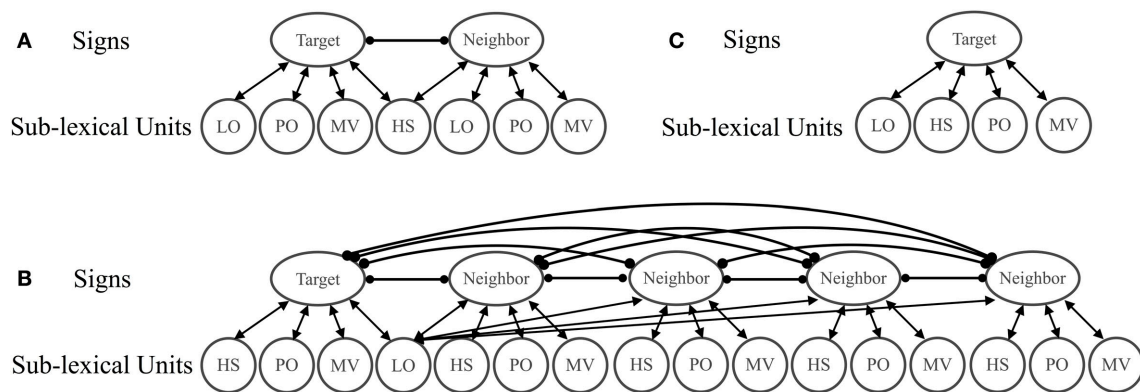
FIGURE 5 | Architecture of sign perception when Neighbor Type was manipulated by varying the number of neighbors. Activation of the target with a handshape neighbor **(A)** or location neighbor **(B)** was compared to activation of the target without a neighbor **(C)**. Neighbors were considered facilitatory if the target item with a neighbor **(A,B)** became active more quickly than the target item without a neighbor **(C)**.
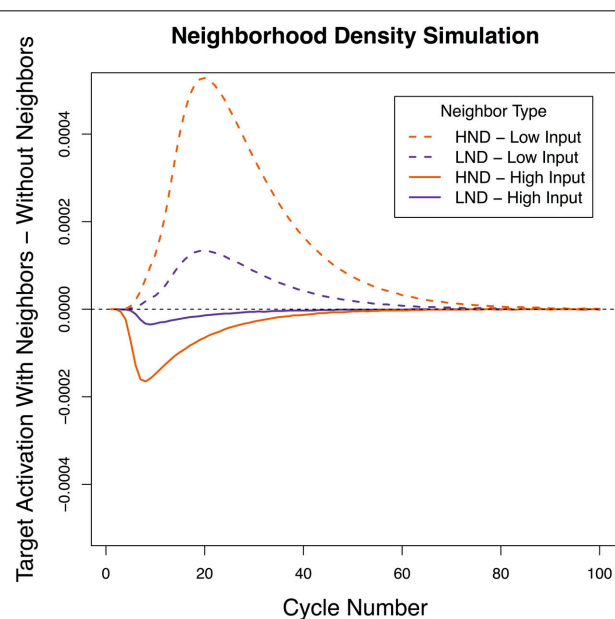

FIGURE 6 | Net contribution of a handshape neighbor and a location neighbor when the number of neighbors was manipulated. Both high and low levels of external input are presented. Both handshape and location neighbors had a net facilitatory effect on the target when the external input was low, and both handshape and location neighbors had a net inhibitory effect on the target when the external input was high.

more inhibitory. These results suggest that the pattern of reversals linked to location and handshape in sign recognition cannot be reduced to differences in neighborhood density, a lexical property. We will discuss this pattern in more depth in the General Discussion.

GENERAL DISCUSSION

The aim of the present study was to computationally test the hypothesis that behavioral patterns in sign recognition can be accounted for using the same lexical access mechanisms that have been proposed for spoken language. Specifically, we investigated whether the opposing effects observed for location and handshape can be obtained in a lexical network that employs universal (language-general) activation principles mediated by language-specific facts about activation levels and neighborhoods.

To do so, we created a spreading activation network with two levels of representation (sub-lexical and lexical) and two types of activation: facilitatory, bidirectional connections between sub-lexical and lexical units; and inhibitory, activation-scaled, unidirectional connections between lexical units (Chen and Mirman, 2012). We then systematically varied three relatively peripheral facts about this network: (1) the timing with which sub-lexical units become active during perception, (2) the resting activation of the sub-lexical units, and (3) the number of lexical neighbors of a target sign. These factors were orthogonally tested in a simulated recognition task with parameters drawn from empirical data about sign languages [specifically: (1) the timing of the perception of location vs. handshape, (2) the sub-lexical frequency of locations vs. handshapes, and (3) the number of a target's location neighbors vs. handshape neighbors].

We found that the specific pattern of facilitation and inhibition reported in sign recognition was obtained when the timing of sub-lexical activation (Simulation 1) and the level of sub-lexical resting activation (Simulation 2) were varied in a manner consistent with real-world facts about location and handshape. We were not able obtain the observed pattern of results when the number of lexical neighbors was similarly varied (Simulation 3). Before drawing conclusions from these results, we wish to address why the network presented a different pattern of results depending on whether sub-lexical or lexical properties were manipulated.

To understand why variations in properties of the shared sub-lexical unit (timing/resting activation) determined whether the net contribution of the neighbor was facilitatory or inhibitory but variations in the size of the lexical neighborhood did not, it is useful to return to the basic principle at the heart of Chen and Mirman (2012)'s architecture: strong neighbors inhibit target processing while weak neighbors facilitate processing. Differences

in the timing and resting activation of a shared sub-lexical unit directly influence how active the neighbor becomes, which in the Chen and Mirman architecture determines whether its net contribution to the target will be negative or positive. In other words, variation in the sub-lexical properties can change the *polarity* of the activation flowing to the target from net positive to net negative. This is why the sub-lexical variations we explored in Simulations 1 and 2 led to differing patterns of facilitation and inhibition. What, then, is the effect of giving a target sign fewer or more neighbors, as in Simulation 3? The crucial fact in this case is that varying the number of neighbors a target has does not influence whether the neighbors themselves are strongly or weakly activated. Because all the neighbors in this model are activated by the same sub-lexical unit, the amount of activation they receive is the same. Therefore, whatever the effect of a single neighbor is in this model, the effect of multiple neighbors will be the same. While the neighbors will become more strongly or less strongly active based on the properties of the sub-lexical units, all of the target item's neighbors will either be net facilitatory or net inhibitory but not both. In other words, the number of neighbors thus does not change the *polarity* of the activation flowing to the target but it does influence the *magnitude*.

In this paper, we attempted to simulate a set of experimental data in order to test the theory that lexical access is accomplished by the same mechanisms in signed and spoken language. Our interpretations about the theory instantiated by the simulation necessarily depend on the assumptions made both in the creation of the simulation and in the design of the original experiments. One concern is that the definition of neighbors used by Carreiras et al. (2008) differs from what is used in research on spoken language. At the moment it is unclear which definition is most appropriate for sign processing (and across different ages of acquisition: Mayberry and Witcher, 2005) and more work is needed to decide this issue. We note, however, that the one-feature-shared definition may have more generalizability than the all-but-one-shared definition simply as there are very few minimal pairs in sign languages relative to spoken languages (van der Kooij, 2002). In addition, the behavioral data modeled here was from LSE signers. More work is needed to explore the generalizability of these results across signed languages. Lastly, the behavioral data modeled in this study consisted of only 4 datapoints from LSE: average reaction times for high vs. low location density and high vs. low handshape density (Carreiras et al., 2008). Likewise, the estimates of sub-lexical frequency and neighborhood density were also based on averages rather than particular lexical items. Future behavioral and computational work is needed to test the model using item-level (and ideally, trial-level) reaction times, sub-lexical frequency and neighborhood density estimates, and timing estimates (e.g., Balota et al., 2007), as well as to measure the goodness of fit of the model. As it stands, this work serves as a proof of concept that the same mechanism for lexical access could underlie both sign and word perception.

The goal of the work presented here was to examine a particular pattern of behavior in lexical access using a set of tightly controlled simulations. In the same way that laboratory experiments make it possible to test the effects of a small set of variables

in isolation, this approach made it possible to orthogonally test the effects of neighborhood density, sub-lexical frequency, and timing. The downside of controlling simulations or experiments so tightly is that it reduces ecological validity. In humans, a number of factors—lexical familiarity (Carreiras et al., 2008) and other neighbor types (Corina and Hildebrandt, 2002; Mayberry and Witcher, 2005; Corina and Knapp, 2006a; Dye and Shih, 2006) to name two—in addition to those modeled here play a role in lexical access. We see computational modeling as an exciting tool to understand sign processing, and hope that over time models like the one presented here can be elaborated to account for many of these factors.

With these assumptions in mind, these results suggest that the pattern of reversals in sign recognition arise because of variation in the activation of sub-lexical units rather than lexical units. In particular, our simulations are consistent with the idea that the sub-lexical feature of location is more robustly encoded or activated earlier than handshape (leading to greater neighbor activation). This prediction connects nicely with other behavioral results. As was mentioned in the introduction, location is misperceived less frequently (Orfanidou et al., 2009), remembered more easily (Thompson et al., 2005), and is produced more accurately by aphasic (Corina, 2000) and unimpaired individuals (Klima and Bellugi, 1979; Hohenberger et al., 2002) than other sub-lexical features. Since activation level correlates with accuracy in spreading activation networks, these empirical results are compatible with our proposal that location representations are able to accrue more activation than handshape representations. More empirical research attempting to elucidate the locus of these various effects is certainly needed.

Our success in modeling the effects of location and handshape in Simulations 1 and 2 provides evidence that there may be universal principles governing the way the mental lexicon is accessed. Even though location and handshape are elements that are unique to sign languages, it appears that their influence on recognition can be modeled using the same principles that have been used to explain lexical access across tasks in spoken and written language. We wish to note that our results do not rule out the possibility that there are sign language-specific factors that influence lexical processing (e.g., distinct "what" vs. "where" processing streams in visual perception). They do, however, indicate that such factors are not necessary to account for the empirical data on reversals. Our investigation suggests that—like the commonalities observed in the grammars of signed and spoken languages—the mind stores and accesses words in the same manner, no matter the modality (spoken, print, or signed).

REFERENCES

Andrews, S. (1992). Frequency and neighborhood effects on lexical access: lexical similarity ororthographic redundancy? *J. Exp. Psychol. Learn. Mem. Cogn.* 18, 234. doi: 10.1037/0278-7393.18.2.234

Balota, D. A., Yap, M. J., Hutchison, K. A., Cortese, M. J., Kessler, B., Loftis, B., et al. (2007). The English lexicon project. *Behav. Res. Methods* 39, 445–459. doi: 10.3758/BF03193014

Baus, C., Gutiérrez-Sigut, E., Quer, J., and Carreiras, M. (2008). Lexical access in Catalan Signed Language (LSC) production. *Cognition* 108, 856–865. doi: 10.1016/j.cognition.2008.05.012

Berent, I., Dupuis, A., and Brentari, D. (2013). Amodal aspects of linguistic design. *PLoS ONE* 8:e60617. doi: 10.1371/journal.pone.0060617

Brentari, D. (1998). *Prosodic Model of Sign Language Phonology*. Cambridge, MA: MIT Press.

Carreiras, M. (2010). Sign language processing. *Lang. Linguist. Compass* 4, 430–444. doi: 10.1111/j.1749-818X.2010.00192.x

Carreiras, M., Gutiérrez-Sigut, E., Baquero, S., and Corina, D. (2008). Lexical processing in Spanish Sign Language (LSE). *J. Mem. Lang.* 58, 100–122. doi: 10.1016/j.jml.2007.05.004

Chen, Q., and Mirman, D. (2012). Competition and cooperation among similar representations: toward a unified account of facilitative and inhibitory effects of lexical neighbors. *Psychol. Rev.* 119, 417. doi: 10.1037/a0027175

Coltheart, M., Davelaar, E., Jonasson, J. T., and Besner, D. (1977). "Access to the internal lexicon," in *Attention and Performance VI*, ed S. Dornic (Hillsdale, NJ: Lawrence Erlbaum Associates).

Corina, D. (2000). "Some observations regarding paraphasia in American Sign Language," in *The Signs of Language Revisited: an Anthology to Honor Ursula Bellugi and Edward Klima*, eds K. Emmorey and H. Lane (Mawah, NJ: Lawrence Erlbaum Associates Inc.), 493–507.

Corina, D. P., and Emmorey, K. (1993). "Lexical priming in American sign language," in *Poster Presented at the 34th Annual Meeting of the Psychonomics Society* (Washington, DC).

Corina, D. P., and Hildebrandt, U. C. (2002). "Psycholinguistic investigations of phonological structure in ASL," in *Modality and Structure in Signed and Spoken Languages*, eds R. Meier, K. Cormier, and D. Quinto-Pozos (Cambridge: Cambridge University Press), 88–111.

Corina, D. P., and Knapp, H. (2006a). "Lexical retrieval in american sign language production. In papers in laboratory phonology," in *Varieties of Phonological Competence*, eds L. M. Goldstein, D. H. Whalen, and C. T. Best (Berlin: Mouton de Gruyter), 213–240.

Corina, D. P., and Knapp, H. (2006b). "Psycholinguistic and neurolinguistic perspectives on sign language," in *Handbook of Psycholinguistics*, eds M. Traxler and M. Gernsbacher (London: Elsevier), 1001–1024.

Dahan, D., Magnuson, J. S., and Tanenhaus, M. K. (2001). Time course of frequency effects in spoken-word recognition: evidence from eye movements. *Cogn. Psychol.* 42, 317–367. doi: 10.1006/cogp.2001.0750

Davis, C. J., Perea, M., and Acha, J. (2009). Re(de)fining the orthographic neighborhood: the role of addition and deletion neighbors in lexical decision and reading. *J. Exp. Psychol. Hum. Percept. Perform.* 35, 1550–1570. doi: 10.1037/a0014253

Dell, G., and Gordon, J. (2003). "Neighbors in the lexicon: friend or foe?," in *Phonetics and Phonology in Language Comprehension and Production*, eds N. O. Schiller and A. S. Meyer (Berlin: Walter de Gruyter), 9–37.

Dell, G. S. (1986). A spreading activation theory of retrieval in language production. *Psychol. Rev.* 93, 283–321. doi: 10.1037/0033-295X.93.3.283

Dufour, S., and Peereman, R. (2003). Inhibitory priming effects in auditory word recognition: when the target's competitors conflict with the prime word. *Cognition* 88, B33–B44. doi: 10.1016/S0010-0277(03)00046-5

Dupuis, A., and Berent, I. (2013). "Lexical Access to Signs is Automatic," in *Paper Presented at the Theoretical Issues in Sign Language Research Conference 11* (London).

Dye, M. W. G., and Shih, S. (2006). "Phonological priming in British Sign Language," in *Papers in Laboratory of Phonology*, Vol. 8, eds L. M. Goldstein, D. H. Whalen, and C. T. Best (Berlin: Mouton de Gruyter), 243–263.

Emmorey, K. (2002). *Language, Cognition, and the Brain: Insights from Sign Language Research*. Mahwah, NJ: Lawrence Erlbaum Associates.

Emmorey, K., and Corina, D. (1990). Lexical recognition in sign language: effects of phonetic structure and morphology. *Percept. Mot. Skills* 71, 1227–1252. doi: 10.2466/PMS.71.7.1227-1252

Emmorey, K., McCullough, S., and Brentari, D. (2010). Categorical perception in american sign language. *Lang. Cogn. Process.* 18, 21–45. doi: 10.1080/01690960143000416

Emmorey, K., Mehta, S., and Grabowski, T. J. (2007). The neural correlates of sign versus word production. *Neuroimage* 36, 202–208. doi: 10.1016/j.neuroimage.2007.02.04

Goldberg, A. M., and Rapp, B. (2008). Is compound chaining the serial-order mechanism of spelling? a simple recurrent network investigation. *Cogn. Neuropsychol.* 25, 218–255. doi: 10.1080/02643290701862332

Goldinger, S. D., Luce, P. A., and Pisoni, D. B. (1989). Priming lexical neighbors of spoken words: effects of competition and inhibition. *J. Mem. Lang.* 28, 501–518. doi: 10.1016/0749-596X(89)90009-0

Goldin-Meadow, S., Mylander, C., and Butcher, C. (1995). The resilience of combinatorial structure at the word level: morphology in self-styled gesture systems. *Cognition* 56, 195–262. doi: 10.1016/0010-0277(95)00662-I

Grainger, J., O'Regan, J. K., Jacobs, A. M., and Segui, J. (1989). On the role of competing word units in visual word recognition: the neighborhood frequency effect. *Percept. Psychophys.* 45, 189–195. doi: 10.3758/BF03210696

Grosjean, F. (1981). Sign and word recognition: a first comparison. *Sign Lang. Stud.* 32, 195–220. doi: 10.1353/sls.1982.0003

Hohenberger, A., Happ, D., and Leuninger, H. (2002). "Modality-dependent aspects of sign language production: evidence from slips of the hands and their repairs in german sign language," in *Modality and Structure in Signed and Spoken Language*, eds R. P. Meier, K. Cormier, and D. Quinto-Pozos (Cambridge: Cambridge University Press), 112–142.

Jackendoff, R. (2002). *Foundations of Language: Brain, Meaning, Grammar, Evolution*. Oxford: Oxford University Press.

Jackendoff, R. (2012). *User's Guide to Thought and Meaning*. New York, NY: Oxford University Press.

Klima, E. S., and Bellugi, U. (1979). *The Signs of Language*. Cambridge, MA: Harvard University Press.

Knobel, M., Finkbeiner, M., and Caramazza, A. (2008). The many places of frequency: evidence for a novel locus of the lexical frequency effect in word production. *Cogn. Neuropsychol.* 25, 256–286. doi: 10.1080/02643290701502425

Lucas, C., and Valli, C. (1992). *Linguistics of American Sign Language: A Resource Text for ASL Users*. Washington, DC: Gallaudet University Press.

Luce, P. A., and Pisoni, D. B. (1998). Recognizing spoken words: the neighborhood activation model. *Ear Hear.* 19, 1. doi: 10.1097/00003446-199802000-00001

MacSweeney, M., Waters, D., Brammer, M. J., Woll, B., and Goswami, U. (2008). Phonological processing in deaf signers and the impact of age of first language acquisition. *Neuroimage* 40, 1369–1379. doi: 10.1016/j.neuroimage.2007.12.047

Marentette, P. F., and Mayberry, R. I. (2000). "Principles for an emerging phonological system: a case study of early ASL acquisition," in *Language Acquisition by Eye*, eds C. Chamberlain, J. Morford, and R. Mayberry (New Jersey, NJ: Lawrence Elrbaum Associates), 71–90.

Mayberry, R. I., and Witcher, P. (2005). "What age of acquisition effects reveal about the nature of phonological processing," in *Center for Research on Language Technical Report*, Vol. 17 (San Diego, CA).

McClelland, J. L. (2009). *Teaching Cognitive Science with MATLAB at Stanford University*. MATLAB Digest.

McClelland, J. L., and Elman, J. L. (1986). The TRACE model of speech perception. *Cogn. Psychol.* 18, 1–86. doi: 10.1016/0010-0285(86)90015-0

McCloskey, M. (1991). Networks and theories: the place of connectionism in cognitive science. *Psychol. Sci.* 2, 387–395. doi: 10.1111/j.1467 9280.1991.tb00173.x

Mirman, D., and Graziano, K. M. (2013). The neural basis of inhibitory effects of semantic and phonological neighbors in spoken word production. *J. Cogn. Neurosci.* 25, 1504–1516. doi: 10.1162/jocn_a_00408

Mirman, D., Kittredge, A. K., and Dell, G. S. (2010). "Effects of near and distant phonological neighbors on picture naming," in *Proceedings of the 32nd Annual Conference of the Cognitive Science Society* (Austin, TX: Cognitive Science Society), 1447–1452.

Mirman, D., and Magnuson, J. S. (2008). Attractor dynamics and semantic neighborhood density: processing is slowed by near neighbors and speeded by distant neighbors. *J. Exp. Psychol. Learn. Mem. Cogn.* 34, 65. doi: 10.1037/0278-7393.34.1.65

Mirman, D., McClelland, J. L., Holt, L. L., and Magnuson, J. S. (2008). Effects of attention on the strength of lexical influences on speech perception: behavioral

experiments and computational mechanisms. *Cogn. Sci.* 32, 398–417. doi: 10.1080/03640210701864063

Mishkin, M., Ungerleider, L. G., and Macko, K. A. (1983). Object vision and spatial vision: two cortical pathways. *Trends Neurosci.* 6, 414–417. doi: 10.1016/0166-2236(83)90190-X

Morford, J. P., and Carlson, M. L. (2011). Sign perception and recognition in non-native signers of ASL. *Lang. Learn. Dev.* 7, 149–168. doi: 10.1080/15475441.2011.543393

Orfanidou, E., Adam, R., McQueen, J. M., and Morgan, G. (2009). Making sense of nonsense in British Sign Language (BSL): the contribution of different phonological parameters to sign recognition. *Mem. Cogn.* 37, 302 3315. doi: 10.3758/MC.37.3.302

Peschl, M. F., and Scheutz, M. (2001). "Explicating the epistemological role of simulation in the development of theories of cognition," in *Proceedings of the 7th International Colloquium on Cognitive Science (ICCS-01)* (San Sebastian), 274–281.

Petitto, L. A., Zatorre, R. J., Gauna, K., Nikelski, E. J., Dostie, D., and Evans, A. C. (2000). Speech like cerebral activity in profoundly deaf people processing signed languages: implications for the neural basis of human language. *Proc. Nati. Acad. Sci.U.S.A.* 97, 13961–13966. doi: 10.1073/pnas.97.25.13961

Pinker, S. (1994). *The Language Instinct: How the Mind Creates Language.* New York, NY: William Morrow.

Poizner, H., Klima, E. S., and Bellugi, U. (1987). *What the Hands Reveal About the Brain.* Cambridge: MIT Press.

Sadat, J., Martin, C. D., Costa, A., and Alario, F. (2014). Reconciling phonological neighbourhood effects in speech production through single trial analysis. *Cogn. Psychol.* 68, 33–58. doi: 10.1016/j.cogpsych.2013.10.001

Sandler, W. (1989). *Phonological Representation of the Sign: Linearity and Nonlinearity in American Sign Language.* Dordrecht: Foris. doi: 10.1515/9783110250473

Sandler, W., and Lillo-Martin, D. (2006). *Sign Language and Linguistic Universals.* New York, NY: Cambridge University Press. doi: 10.1017/CBO9781139163910

Stokoe, W. C. (1972). *Semiotics and Human Sign Languages*, Vol. 21. The Hague: Walter de Gruyter.

Thompson, R., Emmorey, K., and Gollan, T. H. (2005). "Tip of the fingers" experiences by deaf signers insights into the organization of a sign-based lexicon. *Psychol. Sci.* 16, 856–860. doi: 10.1111/j.1467-9280.2005.01626.x

van der Hulst, H. (1993). Units in the analysis of signs. *Phonology* 10, 209–242. doi: 10.1017/S095267570000004X

van der Kooij, E. (2002). *Phonological Categories in Sign Language of The Netherlands: The Role of Phonetic Implementation and Iconicity.* Utrecht: LOT.

Vitevitch, M. S. (1997). The neighborhood characteristics of malapropisms. *Lang. Speech* 40, 211–228. doi: 10.1121/1.415242

Vitevitch, M. S. (2002). The influence of phonological similarity neighborhoods on speech production. *J. Exp. Psychol. Learn. Mem. Cogn.* 28, 735. doi: 10.1037//0278-7393.28.4.735

Wilbur, R. B. (1979). *American Sign Language and Sign Systems: Research and Application.* Baltimore, MD: University Park Press.

Phonological reduplication in sign language: Rules rule

Iris Berent[1]*, Amanda Dupuis[1] and Diane Brentari[2]

[1] Department of Psychology, Northeastern University, Boston, MA, USA
[2] Department of Linguistics, University of Chicago, Chicago, IL, USA

Productivity—the hallmark of linguistic competence—is typically attributed to algebraic rules that support broad generalizations. Past research on spoken language has documented such generalizations in both adults and infants. But whether algebraic rules form part of the linguistic competence of signers remains unknown. To address this question, here we gauge the generalization afforded by American Sign Language (ASL). As a case study, we examine reduplication (X→XX)—a rule that, *inter alia*, generates ASL nouns from verbs. If signers encode this rule, then they should freely extend it to novel syllables, including ones with features that are unattested in ASL. And since reduplicated disyllables are preferred in ASL, such a rule should favor novel reduplicated signs. Novel reduplicated signs should thus be preferred to nonreduplicative controls (in rating), and consequently, such stimuli should also be harder to classify as nonsigns (in the lexical decision task). The results of four experiments support this prediction. These findings suggest that the phonological knowledge of signers includes powerful algebraic rules. The convergence between these conclusions and previous evidence for phonological rules in spoken language suggests that the architecture of the phonological mind is partly amodal.

Keywords: phonology, sign langauge, rules, reduplication, lexical decision

INTRODUCTION

Productivity is the hallmark of linguistic competence (Chomsky, 1957). English speakers, for instance, routinely extend their linguistic knowledge to novel forms that they have never heard before (e.g., *blogs, emails, sms's*). For generative theories of language, such generalizations immediately suggest that the language faculty encodes abstract algebraic rules (Chomsky and Halle, 1968; Chomsky, 1980; Fodor and Pylyshyn, 1988; Pinker and Prince, 1988; Prince and Smolensky, 1993/2004; Pinker, 1994). But whether rules exist, and whether they are linguistic is a matter of debate.

Dozens of connectionist models have shown that linguistic generalizations can emerge in systems that lack rules altogether (Rumelhart and McClelland, 1986; Elman, 1993; Elman et al., 1996; Seidenberg and Jeffery, 1999; McClelland and Patterson, 2002; Haskell et al., 2003; Bybee and McClelland, 2005; Elman, 2005; Bybee, 2008; McClelland, 2009; McClelland et al., 2010; Ramscar and Dye, 2011). Moreover, all previous attempts to adjudicate between rule- and associative-based accounts have been so far limited to spoken language (e.g., Rumelhart and McClelland, 1986; Fodor and Pylyshyn, 1988; Pinker and Prince, 1988; Marcus, 1998). This lacuna raises the question of whether algebraic rules—if they exist—are specific to spoken communication, or whether they form part of the language faculty, generally.

To address these questions, the present research gauges the role of algebraic rules in American Sign Language (ASL). We begin by considering what algebraic rules are, and how they differ from competing (nonalgebraic) associative mechanisms. We next outline how one can adjudicate between these rival accounts by systematically probing the scope of linguistic generalizations.

We first evaluate this question in light of computational and experimental results from spoken languages. These conclusions set the stage for our investigation of rules in sign language.

COMPETING ACCOUNTS OF LINGUISTIC GENERALIZATIONS: RULES vs. ASSOCIATIONS

To appreciate the anatomy of a rule, let us begin by considering the English plural formation rule as a case study (Pinker, 1999). The plural rule generates plural forms by copying the singular noun stem (N_{stem}) and appending the suffix s to its end ($N_{stem} + s$). This simple description entails several critical assumptions concerning mental architecture (Fodor and Pylyshyn, 1988; Pinker and Prince, 1988; Marcus, 2001; for a glossary, see **Box 1**). First, it assumes that the mind encodes ***abstract categories*** (e.g., noun stem, N_{stem}), and such categories are distinct from their instances (e.g., *dog, letter*). Second, mental categories are potentially ***open-ended***—they include not only familiar instances (e.g., the familiar nouns *dog, cat*) but also novel ones. Third, within such category, all instances—familiar or novel—are equal members of this class. Thus, mental categories form ***equivalence classes***. Fourth, mental processes manipulate such abstract categories—in the present case, it is assumed that the plural rule copies the N_{stem} category. Doing so requires that ***rules operate on algebraic variables***, akin to variables from algebraic numeric operations (e.g., $X \rightarrow X+1$)[1]. Finally, because rule description appeals only to this abstract category, the rule will

---

[1] Algebraic rules, as discussed here, are distinct from the technical definition of linguistic rules (mappings from inputs to outputs)—a notion that contrasts with constraints (operations over outputs). Indeed, linguistic rules and constraints both apply to structured representations by virtue of their constituent

---

Box 1 | Glossary.

- **Productivity**. The capacity to extend linguistic generalizations to novel instances.
- **Across the board generalizations**. Generalizations that extend to any member of a class, actual or potential, regardless of their familiarity or similarity to familiar instances.
- **Algebraic rules**. Mental operations that can potentially extend regularities across the board. These generalizations are supported by various representational capacities, including the capacity (a) to form equivalence classes; (b) to operate on entire classes using variables; and (c) to distinguish types (Noun) from individual tokens (*dog*).
- **Equivalence class**. A class of elements whose members (actual or potential) are all treated alike with respect to a given generalization. For example, the English plural formation rule (Noun$_{stem}$+s) treats all "noun stems" alike—it applies to either familiar English stems (e.g., *dog*) or novel ones that are nonnative to English (e.g., *ch* in *Chanukah*).

---

apply equally to any of its members, irrespective of whether any given member is familiar or novel, and regardless of its similarity to existing familiar items[2]. As a result, algebraic rules potentially extend to any member of a class—a property known as ***across-the-board generalizations***.

The hypothesis that the language system encodes algebraic rules is consistent with myriad of linguistic data, showing that speakers of many languages extend their knowledge to novel forms. Generalization, however, does not, in and of itself, demonstrate that the mind encodes rules. Indeed, connectionist networks have been shown to exhibit generalizations despite the elimination of algebraic mechanisms—they encode no abstract categories (e.g., Noun) distinct from their instances (e.g., *dog*), and consequently, they lack mechanisms that operate on entire classes (i.e., operations over variables). Generalizations in such models (e.g., to *rogs*) depend not on variables standing for abstract classes (N$_{stem}$ +s), but rather on the association between their specific instances (e.g., between *rog-rogs* and *dog-dogs*); the mechanisms that produce regular forms (e.g., *rats*) are indistinguishable from the ones responsible for the formation of exceptions (e.g., *mice*). Yet, such models have been shown to capture significant aspects of speakers' knowledge of existing forms, and even generalize to novel ones (Rumelhart and McClelland, 1986; Elman et al., 1996; McClelland and Plaut, 1999; McClelland and Patterson, 2002).

Given that rules and associations can both lead to generalizations, merely showing that people can generalize linguistic functions cannot adjudicate between competing accounts of language. Nonetheless, algebraic and associationist accounts are not homologous. Their differences become evident once we take a closer look at the scope of generalizations.

COMPUTATIONAL TESTS OF COMPETING ARCHITECTURES: THE SCOPE OF LINGUISTIC GENERALIZATIONS

Algebraic and associationist architectures can both generalize, but the generalizations they attain differ in scope. The evidence

---

structure, and as such, they both invoke the same notion of algebraic rules examined here.

[2]The potential of a rule to apply across the board does not mean that the operation of the rule is never circumvented. The English plural rule, for instance, is blocked whenever irregular counterexamples are retrieved from memory. Such limitations, however, are imposed by factors external to the rule (e.g., conflicting rules, lexical stipulations), rather than from limitations on the inherent capacity of the rule to generalize, and as such, they are irrelevant to evaluating the scope of potential generalizations.

comes from computational simulations that systematically gauge the scope of generalizations of a reduplication rule—a function that is commonly found in the morpho-phonology of many languages (e.g., McCarthy, 1986; Suzuki, 1998), and forms the center of our following investigation of sign language. In its simplest form, the reduplication function (X→XX) copies some prosodic unit X (e.g., a syllable; e.g., *baba, dada, tata*). Our question here is what kind of computational system—algebraic or associative—is necessary to freely generalize the reduplicative function to any class member.

Rules, by definition, generalize across the board, so such generalizations are clearly consistent with an algebraic system. A series of simulations by Gary Marcus suggests that they are inconsistent with (nonalgebraic) connectionist networks (Marcus, 1998, 2001). This is not because connectionist networks are categorically unable to generalize; Marcus showed that the reduplication function is successfully learnable by various connectionist networks (feed-forward and simple recurrent networks). But unlike symbolic architectures, generalizations in these networks are systematically limited by the similarity of novel test items to familiar instances.

Novel test items that shared all their features with training instances (i.e., generalizations within their training space) yielded robust generalizations. But when presented with test items including unfamiliar features (i.e., items falling outside the training space), the networks failed to generalize the reduplication function. For example, a network trained on reduplicants with a labial feature (e.g., *papa, mama*) might readily generalize to a novel labial *baba*, as the network can exploit the association between the two labial features in the training items. But since this generalization is solely based on feature-association in training items (e.g., the labial-labial feature), once presented with a velar test item (e.g., *gaga*), generalization will likely fail, as the model lacks knowledge relevant to the reduplication of the velar feature. Subsequent work showed that, absent algebraic rules, the failure to generalize to dissimilar novel items also emerges in the Maximum Entropy Model (Berent et al., 2002)—an influential computational account of phonology (Hayes and Wilson, 2008). Thus, models that lack algebraic mechanisms can generalize, but they cannot do so systematically, across the board.

THE SCOPE OF PHONOLOGICAL GENERALIZATIONS IN SPOKEN LANGUAGE

The systematic links between the architecture of a computational system and its capacity to generalize are significant because they

can be used to gauge the architecture of the language system. If the language faculty encodes algebraic rules, then people should extend generalizations across the board, but if they rely on associations, then generalizations will apply only to novel items that share their features with familiar linguistic exemplars.

Previous work on spoken language has tested this prediction using the reduplication function. The evidence comes from speakers of Hebrew—a language that (like other Semitic languages) systematically restricts the location of reduplicated elements in its stems. Hebrew allows identical consonants to occur at the right edge of the stem (e.g., *salal*, "paved"), but bans them in its beginning (e.g., *lalas*; Greenberg, 1950; Leben, 1973; McCarthy, 1986, 1989). Thus, XYY stems (X, Y = any consonant) are well-formed whereas XXY stems are ill-formed.

A large body of experimental research shows that Hebrew speakers generalize this restriction to novel forms (Berent and Shimron, 1997; Berent et al., 2001a,b, 2002, 2004, 2006, 2011, 2012a; Berent and Shimron, 2003)—a conclusion that converges with artificial language experiments with adults (Endress et al., 2005; Toro et al., 2008) and infants (Marcus et al., 1999, 2007; Gervain et al., 2008, 2012). Such results demonstrate that the reduplication function is productive, but they do not attest to the scope of the generalization, and consequently, they do not distinguish between rule-based and associative explanations. Specifically, a generalization to a novel form (e.g., *tagag*) can either occur because Hebrew speakers encode the reduplicative structure of this stem (i.e., as YX_iX_{Ci}, where the Ci is a copy of element i) or because they associate it with existing stems (e.g., *xagag*, "he celebrated").

To adjudicate between these competing accounts, one can examine whether Hebrew speakers generalize the identity function to novel stems whose phonemes and features are unattested in Hebrew. For example, Hebrew lacks the phoneme corresponding to the English *th* (e.g., *thing*), and its place of articulation (the wide value of the tongue tip constriction area feature, Gafos, 1999) is likewise unattested. Of interest is whether Hebrew speakers favor novel well-formed YXX stems like *kathath* to their XXY counterparts (e.g., *thathak*, Berent et al., 2002). Findings from a series of experiments suggest that they do just that. Specifically, *thathak*-type forms are less acceptable in rating experiments, and since such ill-formed items are less word-like, they are also classified as nonwords more readily in lexical decision.

The results concerning the reduplication rule are particularly significant because reduplication (and its mirror image, identity restrictions) is fundamental to many phonological and morphological systems (Suzuki, 1998; Frampton, 2009). Accordingly, finding that people extend the reduplication rule across the board suggests that the phonological system of *spoken* language exhibits unbounded productivity—a capacity that would put phonological generalizations on par with syntactic rules. Our present research asks whether algebraic rules also form part of sign language.

PHONOLOGICAL GENERALIZATIONS IN SIGN LANGUAGE
Every established sign language exhibits a phonological system of intricate design. As in spoken phonology, signed phonological systems encode the hierarchical organization of discrete distinctive features (Brentari, 1998; Sandler and Lillo-Martin, 2006), they represent the syllable—a prosodic unit that is demonstrably distinct from a morpheme (Brentari, 1998; Sandler and Lillo-Martin, 2006), and constrain their sonority profile (Stokoe, 1960; Klima and Bellugi, 1979; Corina, 1990; Perlmutter, 1992; Brentari, 1993, 1994, 1998; Corina and Sandler, 1993; Brentari, 2006; Sandler and Lillo-Martin, 2006; Sandler, 2008; Jantunen and Takkinen, 2010; Wilbur, 2012). Experimental research on sign languages has further shown that signers—both adults (Lane et al., 1976; Newport, 1982; Hildebrandt and Corina, 2002; Emmorey et al., 2003; Baker et al., 2005; Best et al., 2010) and infants (Baker et al., 2006; Palmer et al., 2012)—encode phonological features as phonetic categories, subject to perceptual narrowing in the first year of life (Baker et al., 2006; Palmer et al., 2012). Moreover, distinct feature classes differ in their contribution to language processing. Location information, specifically, is particularly salient to lexical access (Emmorey and Corina, 1990; Corina and Hildebrandt, 2002; Thompson et al., 2005; Baus et al., 2008; Carreiras et al., 2008; Orfanidou et al., 2009; Gutiérrez et al., 2012); it provides a strong cue for similarity (Hildebrandt and Corina, 2002; Bochner et al., 2011); and it is acquired earlier (Siedlecki and Bonvillian, 1993) and more accurately (Marentette and Mayberry, 2000; Morgan, 2006) during first-language acquisition. Other studies have suggested that typical (Morgan, 2006; Morgan et al., 2007) and disordered (Marshall et al., 2006) acquisition of sign language is constrained by the complexity of features and their distance from the body (Meier, 2000; Meier et al., 2008)—a factor also affecting adult signers (Poizner et al., 1981).

Most of this work, however, has focused on individual phonological features, rather than the restrictions governing their combination, and with a couple of exceptions (Carreiras et al., 2008), most results obtained from existing signs. There is also some evidence that signers are sensitive to phonotactic legality (Orfanidou et al., 2010) and the number of syllables in novel signs (Brentari et al., 2011)—phonological units distinct from morphemes (Berent et al., 2013). Nonetheless, it is uncertain whether such knowledge reflects algebraic rules, or the statistical structure of the lexicon—a factor to which signers are acutely sensitive (Carreiras et al., 2008). Whether signers possess the capacity for unbounded productivity—the hallmark of powerful algebraic mechanisms—is unknown. No previous experimental research has addressed this question.

Only one previous study examined the capacity of 7.5 month-old hearing infants to acquire rules from novel signs (Rabagliati et al., 2012). The results, however, were mixed. While participants in this experiment freely extended the YXX rule, they failed to acquire the XXY regularity—a rule they can readily learn from speech stimuli. Moreover, the (limited) generative mechanisms available to infants might not necessarily form part of the linguistic competence of adult signers. One thus wonders whether algebraic rules are inherent to the phonological mind (Berent, 2013a), generally, or to the speech modality, specifically[3].

---

[3]The algebraic account is further challenged by the iconicity of signs (Ormel et al., 2009; Eccarius and Brentari, 2010; Thompson et al., 2010; Brentari, 2011), which has been shown to affect their on-line processing by adults

OUR PRESENT EXPERIMENTS: DO SIGNERS EXTEND THE REDUPLICATION FUNCTION ACROSS THE BOARD?

Our present study examines the scope of phonological generalizations of the reduplication function. We chose the case of reduplication for two reasons. First, reduplication has been the subject of intense computational effort, so the principled limitations of nonalgebraic mechanisms to extend this function are well documented. Second, reduplication is central to the phonology and morphology of sign language. Like spoken phonological systems, signed phonological systems exhibit various forms of reduplication (Klima and Bellugi, 1979; Sandler and Lillo-Martin, 2006; Wilbur, 2009). One such form generates ASL nouns by reduplicating their verbal counterparts—this process maintains the handshape, location and directionality of movement of their base verb, but invariably changes the frequency and manner of movement to become restrained and repeated (Supalla and Newport, 1978)[4]. While this relationship is systematic (Wilbur, 2009), the class of such verb-noun pairs is rather small, and it is unknown whether it is productive (i.e., whether it generalizes to novel signs). Indeed, related research on reduplication in sign language acquisition (Morgan, 2006) has invoked motor, rather than cognitive factors (Meier et al., 2008). Our following research thus asks whether signers (and nonsigners) extend this rule productively, and whether they do so across the board—regardless of whether the reduplicated feature is attested in their language.

Experiments 1–2 present participants with novel disyllabic signs—either reduplicated or nonreduplicated controls, matched for the first syllable. Using X and Y to represent those two syllables, reduplicated and nonreduplicated signs can be denoted as XX and XY, respectively. These syllables are comprised of native ASL features, and their phonotactic structure is otherwise legal. If signers encode the reduplication rule, then they should favor novel reduplicated signs to their nonreduplicated counterparts. Such preference is expected either because reduplication is grammatically better-formed (i.e., unmarked[5]; McCarthy and Prince, 1995) or because, as a type, reduplicated signs are far more

frequent in ASL than nonreduplicated disyllables. Either way, XX novel signs should appear more "sign-like." Accordingly, novel XX signs should be rated higher than XY controls, and they should be harder to classify as "nonwords" in lexical decision. Experiments 1 (rating) and 2 (lexical decision) address these questions.

The hallmark of algebraic rules, however, is that they support generalizations to *any* member of a class—actual or potential, and past research documented such generalizations in spoken languages. Experiments 3–4 next ask whether unbounded productivity also applies to signs. Experiment 3 elicits ratings of reduplicated signs with unattested handshapes; in Experiment 4, participants perform lexical decision. If reduplication is represented by an algebraic rule, then XX forms should appear more sign-like even when the reduplicated form includes an unattested feature.

PART 1: GENERALIZATION TO ATTESTED FEATURES
EXPERIMENT 1: OFF-LINE RATING

As a preliminary test, Experiment 1 evaluates signers' sensitivity to reduplication using an off-line rating task. In each trial, participants are presented with a pair of video clips featuring novel ASL signs—a reduplicated XX sign and a nonreduplicated XY control—matched to the reduplicated sign for the initial syllable X (see **Figure 1**). Of interest is whether signers favor novel reduplicated signs to XY controls.

To determine whether this preference is modulated by linguistic experience with ASL, we also elicited similar ratings from a group of nonsigners, native English speakers. Convergence between the two groups will suggest that the effect of reduplication solely stems from sources (linguistic or otherwise) that are independent of linguistic experience with ASL; divergence will suggest that the encoding of reduplication is at least partly modulated by linguistic knowledge.

Methods

Participants. Two groups of adult participants took part in the experiment. One group consisted of twelve Deaf signers who were all exposed to ASL by the age of five (three were exposed to ASL from birth, four by the age of two, and the remaining five by the age of five). The second group consisted of twelve English speakers who were not signers of ASL. Eleven of these participants reported no previous exposure to ASL; one participant had a rudimentary knowledge of the ASL alphabet.

Materials. The materials consisted of short video clips, featuring sixteen pairs of novel disyllabic signs. Within each pair, one member was reduplicated (XX), whereas the other member was nonreduplicated (XY). Pair members were matched for the first syllable (X) and they were phonotactically legal in ASL. A complete list of the materials is presented in Supplementary Material.

These materials were video recordings of a native ASL signer. Prior to the recording, the signer practiced the items so that they are signed naturally. Another native ASL signer recorded the

---

(Thompson et al., 2009, 2010), children (Ormel et al., 2009) and infants (Thompson et al., 2012; but see Emmorey et al., 2004; Bosworth and Emmorey, 2010). Iconicity implies that the representation of signs is continuous and analog, not discrete and digital, as required by the algebraic proposal. However, the effects of iconicity are not specific to sign language (for a recent review of spoken language, see Schmidtke et al., 2014). Moreover, the encoding of phonetic and embodied aspects of signed and spoken words does not preclude the existence of a second format of representation that is algebraic, abstract and fully productive (Brentari, 2007; Mahon and Caramazza, 2008; Eccarius and Brentari, 2010).

[4]It has been argued that the targeted syllables are "light" (i.e., syllables with a single movement component, Brentari, 1998).

[5]Marked structures are (a) disfavored as the output of grammatical processes (de Lacy, 2006); and (b) underrepresented in the language (Prince and Smolensky, 1993/2004). Reduplicated signs meet both requirements for (un)markedness. While many grammatical processes produce reduplicated signs (Klima and Bellugi, 1979; Wilbur, 2009), nonreduplicated disyllables are avoided, resulting in their reduction to monosyllables (Sandler and Lillo-Martin, 2006). In addition, nonreduplicated disyllables are systematically underrepresented in the ASL lexicon. Our inspection of an on-line ASL dictionary (ASLpro.com) identified a total count of 1830 disyllables. Of those, the grand majority (69.7%) are fully reduplicated, 20% are partially reduplicated

whereas only 10.27% are nonreduplicated. Given those observations, XX signs are likely less marked than XY ones.
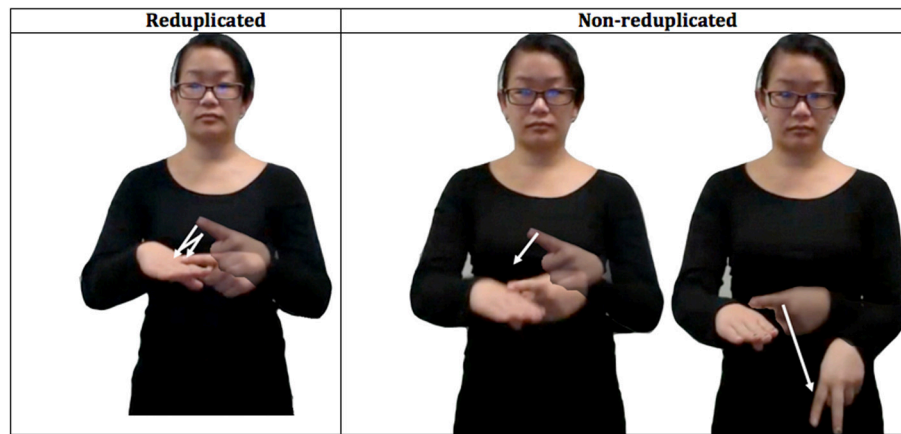
FIGURE 1 | An illustration of the novel signs used in Experiment 1.

instructions to the experiment in ASL. The video recordings of the stimuli were subsequently edited, so that each video clip began immediately upon the initiation of the signing movement and ended with the signer returning to a neutral position. All video clips were inspected for clarity by a fluent ASL signer (DB).

Procedure. In each trial, participants were presented with a matched pair of novel signs (XX and XY, counterbalanced for order). Signers were told that while the stimuli are not ASL signs, they could potentially exist in ASL. Nonsigners received the same instructions, with the added acknowledgement that the task is difficult to perform without knowledge of ASL and the request to "just try to go with your gut feeling." Participants were asked to indicate which pair member is more acceptable as an ASL sign. They were allowed to replay the two options as necessary. Signers were presented with the instructions in ASL, whereas non-signers were presented with English instructions. In this and all experiments, trial order is randomized.

Results and Discussion

Figure 2 plots signers' rating preferences. Results show that on most trials (73%), signers favored reduplicated novel signs to nonreduplicated controls, and these ratings were found statistically different from chance by t-tests [$t1_{(11)} = 5.16$, $p < 0.003$; $t2_{(15)} = 6.04$, $p < 0.0001$]. Nonsigners, by contrast, exhibited no such preference. In fact, nonsigners favored nonreduplicated to reduplicated signs [$M = 33\%$, $t1_{(11)} = -3.68$, $p < 0.006$; $t2_{(15)} = -5.44$, $p < 0.0001$].

Signers' capacity to extract reduplication from novel signs is consistent with the possibility that they rely on an algebraic rule. The contrast between the performance of signers and nonsigners suggests that this rule is informed by their linguistic experience with ASL.

The results from the off-line rating procedure, however, are limited inasmuch as they do not address the role of rules in on-line language processing. To examine this question, we next turn to investigate whether signers might encode the reduplicative structure of signs when a rapid on-line response is required, using the lexical decision task.
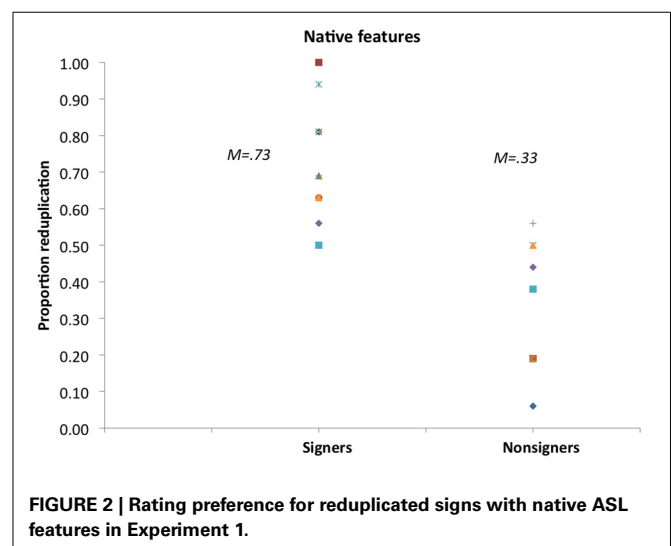
FIGURE 2 | Rating preference for reduplicated signs with native ASL features in Experiment 1.

EXPERIMENT 2: LEXICAL DECISION

Experiment 2 probes signers' sensitivity to reduplication in the lexical decision task. In each trial, participants were presented with a video clip featuring either an attested ASL sign or a novel sign. Within each such category, half of the items exhibited reduplication (XX), whereas the other half was not reduplicated (XY). Participants were asked to quickly determine whether the stimulus is a real ASL sign, and indicate their response by pressing one of two keys (1 = ASL signs; 2 = nonsigns).

If signers can extend the reduplication rule productively, then reduplicated XX signs should be differentiated from nonreduplicated XY controls; and since novel XX signs are grammatically structured and better formed (i.e., unmarked as compared to XY forms), then they should further appear as more sign-like. Consequently, novel XX signs should be harder to classify as nonsigns relative to nonreduplicated XY controls. In contrast, attested ASL signs with reduplication should be classified more readily than their XY counterparts.

Methods

Participants. Participants were the same individuals who took part in the rating experiment (Experiment 1), administered after Experiment 2. Data from one of these participants were excluded from all analyses of Experiment 2 because this individual had reported that he did not understand the task after completing the experiment—an assessment consistent with this individual's accuracy (45%). The results are based on the data of the remaining eleven participants.

Materials. The materials consisted of 16 pairs of ASL signs and 16 pairs of novel ASL signs. Within each category, half of the items were reduplicated, whereas the other half was not reduplicated. The reduplicated ASL signs were all disyllabic nouns that are morphologically related to an ASL verb[6]. The nonreduplicated ASL signs were all ASL compound signs. The reduplicated and nonreduplicated pair members were matched for either hand-shape (in 6/16 pairs) or location (in 10/16 pairs). Novel signs corresponded to the same novel signs used in Experiment 1. All signs (attested ASL and novel) were recorded by the same native signer. The video recordings of the stimuli were subsequently edited, so that each video clip began immediately upon the initiation of the signing movement and ended with the signer returning to a neutral position. All recordings were inspected for clarity by a fluent ASL signer (DB). The complete lists of the novel and existing ASL signs are presented in Supplementary Material.

Procedure. Each trial began with a screen displaying a fixation point. Participants initiated the trial by pressing the spacebar, and their response triggered the presentation of a single video clip (for up to 4 s). Participants were informed that they were about to watch videos of real and novel signs in American Sign Language. They were told that the novel signs are not used in ASL, but they potentially could be "true ASL signs." Participants were asked to determine whether the stimulus was a real ASL sign, and indicate their response by pressing one of two keys (1 = sign, 2 = nonsign). They were instructed to make their response as quickly and as accurately as possible. Slow responses (slower than 2250 ms) triggered the presentation of a warning message (an image of a clock), reminding participants to respond faster. Likewise, participants received computerized feedback on their accuracy (green "smiley" face vs. red "sad" face for correct vs. incorrect responses, respectively).

Prior to the experiment, participants took part in a brief practice session. None of the practice items appeared in the experimental session. In this and all subsequent experiments, response time is reported from the onset of the stimulus.

Results

Outliers (correct responses slower than 3000 ms or faster than 250 ms, less than 1.6% of the total correct responses) were excluded from the analyses of response time. The mean error and

correct response time of signers to ASL signs and novel signs is presented in **Figure 3**.

Errors. An inspection of the error means suggests that signers were sensitive to reduplication. Reduplication elevated errors in response to novel signs, but tended to improve accuracy for existing ASL signs.

These conclusions were supported by the 2 lexicality (sign vs. novel sign) × 2 reduplication (reduplication vs. nonreduplication) ANOVAs, conducted over the error data (arcsine transformed) using both participants (F1) and items (F2) as random variables. The analyses yielded a significant main effect of lexicality $[F1_{(1, 10)} = 8.10$, MSE = 0.058, $p < 0.02$; $F2_{(1, 30)} = 5.48$, MSE = 0.167, $p < 0.03]$ and a marginally significant effect of reduplication $[F1_{(1, 10)} = 2.97$, MSE = 0.027, $p < 0.12$; $F2_{(1, 30)} = 3.11$, MSE = 0.07, $p < 0.09]$. Crucially, the interaction was highly significant $[F1_{(1, 10)} = 15.63$, MSE = 0.065, $p < 0.003$; $F2_{(1, 30)} = 9.28$, MSE = 0.075, $p < 0.005]$[7].

To further probe this interaction, we next tested the effect of reduplication for ASL signs and novel signs, separately. Novel reduplicated signs produced significantly more errors compared to nonreduplicated controls $[t1_{(10)} = 5.68$, $p < 0.0003$; $t2_{(15)} = 4.78$, $p < 0.0003]$. The opposite trend emerged for signs, but it was not significant $[t1_{(10)} = 2.00$, $p < 0.08$; $t2_{(15)} < 1]$.

Response time. **Figure 3** provides the mean correct response time as a function of lexicality and reduplication. The 2 lexicality × 2 reduplication ANOVAs yielded only a reliable main effect of lexicality $[F1_{(1, 10)} = 101.22$, MSE = 6839, $p < 0.00001$; $F2_{(1, 29)} = 45.16$, MSE = 19,450, $p < 0.0001]$ and reduplication $[F1_{(1, 10)} = 29.07$, MSE = 6510, $p < 0.0004$; $F2_{(1, 29)} = 13.63$, MSE = 2216, $p < 0.002]$. The reduplication × lexicality interaction was marginally significant $[F1_{(1, 10)} = 4.77$, MSE = 4949, $p < 0.06$; $F2_{(1, 29)} < 1]$.

Tests of the simple main effect showed that reduplicated signs elicited reliably faster responses compared to nonreduplicated signs $[t1_{(10)} = 9.54$, $p < 0.0001$; $t2_{(15)} = 3.64$, $p < 0.004]$. In contrast, for novel signs, the effect of reduplication was not reliable $[t1_{(10)} = 2.04$, $p < 0.07$; $t2_{(14)} = 1.76$, $p < 0.11]$.

Discussion

Experiment 2 examined whether ASL signers extend the reduplication rule to novel signs. Because reduplicated stimuli are grammatically structured, we expected novel reduplicated signs to appear more sign-like. In accord with this prediction, novel reduplicated signs produced more errors, suggesting that they resemble ASL signs more than nonreduplicated controls. In contrast, for existing ASL signs, reduplication sped up response relative to nonreduplicated controls. These findings demonstrate that participants are sensitive to the reduplicative structure of

---

[6]Some of these ASL signs also have a one movement variant, but these monosyllabic variants were not the ones used in our experiment—all experimental items were invariably disyllables with two full movements.

[7]To ensure that the error results are not due to artifacts associated with binary data, we also submitted the error data to a mixed effects logistic analysis, with lexicality and reduplication as fixed effects (sum-coded) and participants and items as random effects. These analyses yielded a reliable lexicality × reduplication interaction ($\beta = -0.6061$, $SE = 0.122$, $Z = -4.95$, $p < 0.001$).

FIGURE 3 | Lexical decision results for ASL signs and novel signs with native ASL features in Experiment 2. Note: Error bars are 95% confidence intervals for the difference between the means.

FIGURE 4 | An illustration of the four unattested handshapes used in Experiments 3–4.
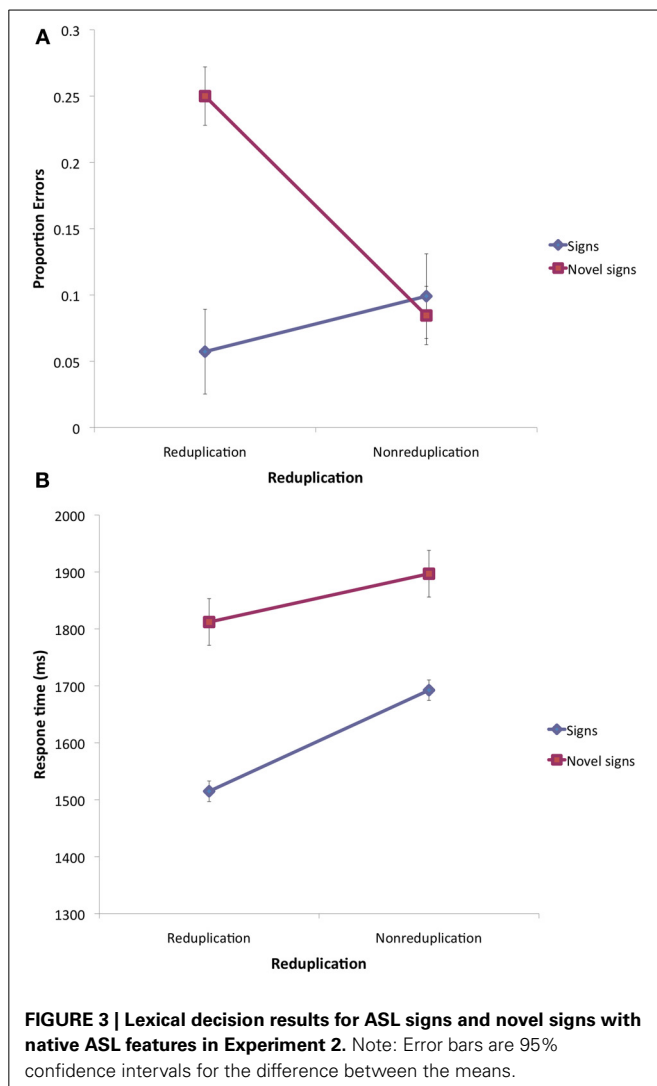
novel signs—an observation consistent with the hypothesis that signers encode productive grammatical rules.

PART 2: GENERALIZATION TO UNATTESTED FEATURES

Experiments 1–2 suggest that signers can extract the reduplication of signs whose features are all native to ASL. The hallmark of algebraic rules, however, is that they support broad generalizations to *any* class member. Accordingly, if signers encode reduplication by a rule (X→XX, where X stands for any syllable), then they should extend it not only to novel syllables with native ASL features (studied in Experiments 1–2) but even to novel syllables with unattested phonological features.

To test this possibility, Experiments 3–4 present participants with novel signs whose reduplicated syllable (X) includes a handshape that is unattested in ASL. Four such handshapes were selected: the OI, EE, V* and the Claw*[8] (see **Figure 4**). These
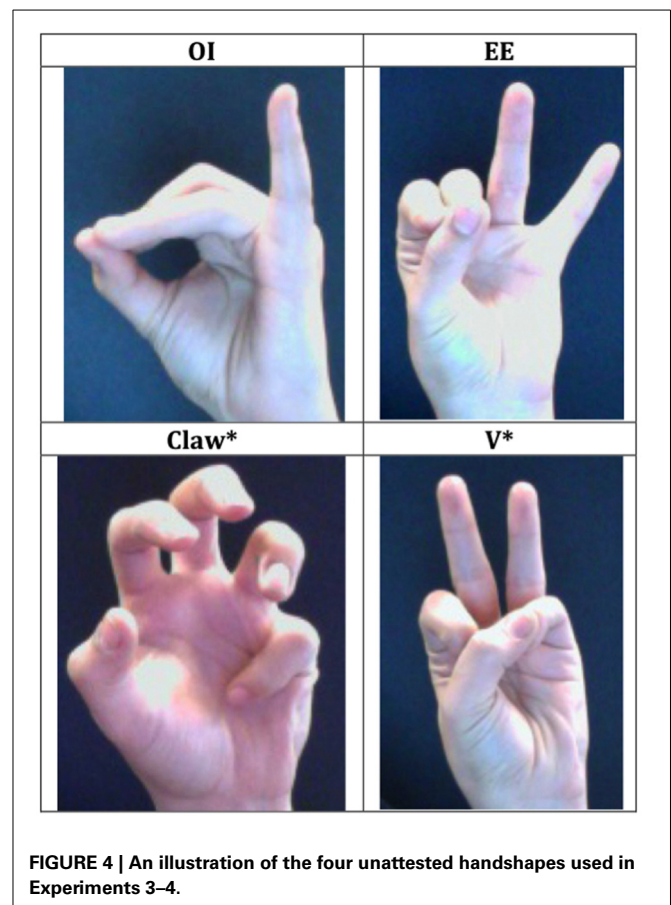
four handshapes are all sign-like, and two of them—the OI and EE handshapes—are attested in Russian Sign Language and Japanese Sign Language. But despite their phonotactic legality, those handshapes are distinctly unattested in ASL, and as such, they are unlikely to readily assimilate to an ASL handshape. This characteristic of the stimuli is significant because past computational results have shown that algebraic rules are necessary to capture the reduplication of unfamiliar features, but they are not indispensable in generalizations to familiar features (Marcus, 1998; Berent et al., 2012b). If participants were to misperceive the unattested handshapes as ASL features, then generalizations to such features would not require reliance on algebraic rules. Our choice of nonnative features was designed to counter this concern.

Each such feature was incorporated in both a reduplicative novel sign (XX) and a nonreduplicative control (XY). In the XY controls, the initial syllable was identical to the reduplicated counterpart (XX), whereas the second syllable Y had a native handshape (see **Figure 5**). Note that the reduplicated signs were statistically less similar to ASL signs, as they included two unattested handshapes—more than in XY controls (with only a single unattested handshape). Accordingly, our experiments pit the contribution of the grammatical reduplication rule against the statistical structure of the ASL lexicon.

Experiment 3 first elicits off-line rating of novel XX and XY signs. To determine whether signers' preferences are informed

by linguistic knowledge, we also obtained similar ratings from a group of English speaking nonsigners. Experiment 4 next examined whether signers extract reduplication on-line, in the lexical decision task.

EXPERIMENT 3: OFF-LINE RATINGS
Methods
Participants were the same twelve Deaf adults and twelve English speakers who took part in Experiment 1 (rating novel ASL signs comprised of native features). Experiment 3 was administered after participants took part in Experiment 1.

Materials. The materials consisted of short video clips, featuring sixteen novel pairs of ASL signs. Within each pair, one member was reduplicated (XX) whereas the other was nonreduplicated (XY), matched to its reduplicated counterpart for the initial syllable (X). In each such member, the syllable X comprised of a handshape that is unattested in ASL, whereas the Y syllable had a native ASL handshape. Four unattested handshapes were used: OI, EE, V* and Claw*. The OI and EE handshapes are attested in Russian Sign Language and Japanese Sign Language; the remaining two handshapes were designed to appear as sign-like. Each such handshape was incorporated in four pairs.

All other features were matched to the novel signs employed in Experiments 1–2. Specifically, each unattested nonsign was created by replacing the handshape in syllable X of the attested nonsigns (used in Experiments 1–2) with one of the four nonnative handshapes mentioned above. Unattested nonsigns matched the attested nonsigns for location, movement, palm-orientation, and handshape in the Y syllable, and these items were thus phonotactically legal in ASL.

All video clips were recorded by a native ASL signer (the same individual featured in all experiments). Prior to the video recording, the signer practiced the signs, to ensure their fluent production. The video clips were subsequently edited, so that each clip began with the initiation of the signing movement and ended with the signer returning to a neutral position. All items were inspected for clarity by a fluent ASL signer (DB).

Procedure. This was identical to Experiment 1.

Results
Figure 6 plots the proportion of trials in which participants favored the reduplicated sign over its nonreduplicated counterpart. An inspection of the means suggests that, on most trials, signers favored the reduplicated signs. T tests, assessing the reliability of this preference across participants' and items' means confirmed that preference for reduplicated signs was reliably different from chance level [$M = 62\%$, $t1_{(11)} = 2.48$, $p < 0.04$; $t2_{(15)} = 2.59$, $p < 0.03$]. In contrast, nonsigners exhibited an opposite preference for nonreduplicated signs [$M = 32\%$, $t1_{(11)} = -2.86$, $p < 0.02$; $t2_{(15)} = -3.81$, $p < 0.002$].

Signers' consistent preference for the reduplicated signs is remarkable given that these stimuli were statistically *less* similar to ASL signs than the nonreduplicative controls. Indeed, XX stimuli included two unattested ASL handshapes (one for each X syllable), whereas XY controls only had one such feature. The consistent preference for reduplication, despite conflicting statistical

information, demonstrates that signers extracted the reduplicative structure. Their capacity to do so with unattested features could imply a productive algebraic rule.

EXPERIMENT 4: LEXICAL DECISION
In Experiment 4, we examine whether signers can extract the reduplication of unattested features in on-line language processing. To this end, we present the same set of novel signs from Experiment 3, mixed with ASL signs (used in Experiment 2) in a lexical decision task. Within each category, half of the stimuli were reduplicated, the others were nonreduplicated. In each trial, participants saw a single stimulus—either an ASL stimulus, or a novel sign with an unattested handshape.

Our experiment addresses two questions. First, we ask whether signers register the presence of unattested features in our materials. If they do, then novel signs with unattested features should be more readily recognized as such. Consequently, lexical decision in Experiment 4 should be faster and more accurate relative to Experiment 2—where the same ASL signs were paired with novel signs whose handshapes are attested in ASL.

Having demonstrated that participants registered the novel handshape faithfully, we can next move to examine our main question—whether signers represent its reduplication. If signers extract the reduplicative structure of novel handshapes, then novel XX signs should appear more sign-like (either because reduplication is less marked, or more frequent in ASL disyllables), hence, they should impair the identification of novel reduplicative signs relative to nonreduplicated controls.

Methods
Twelve Deaf adult, native ASL signers took part in the experiment. These individuals also took part in Experiment 2 prior to completing this experiment. Thus, the order of the four experiments was 2, 4, 1, 3 (i.e., rating and lexical decision for novel signs with attested features, followed by rating and lexical decision of novel signs with unattested features), and they were all administered in a single session. Materials, Instructions and Procedure were the same as in Experiment 2, except that the novel signs had unattested handshapes, as described in Experiment 3. The instructions to the experiment informed participants that they were about to see novel signs that do not occur in ASL, but contain elements that are borrowed from other sign languages.

Results
Do signers register the presence of unattested handshapes?
Before we can examine our main question of interest—whether signers are sensitive to the reduplication of unattested handshapes—we must first establish that signers did in fact register the presence of unattested features in our materials. If they did, then lexical decision should be easier to perform for nonsigns with unattested ASL features (in Experiment 4) compared to those with attested features (in Experiment 2).

To test this possibility, we compared the lexical decision responses in Experiment 4 (with unattested handshapes) to those in Experiment 2 (with attested handshapes) via 2 attestation (attested vs. unattested handshapes) x 2 lexicality (signs vs.

FIGURE 5 | An illustration of the novel signs with unattested handshapes used in Experiments 3, 4.
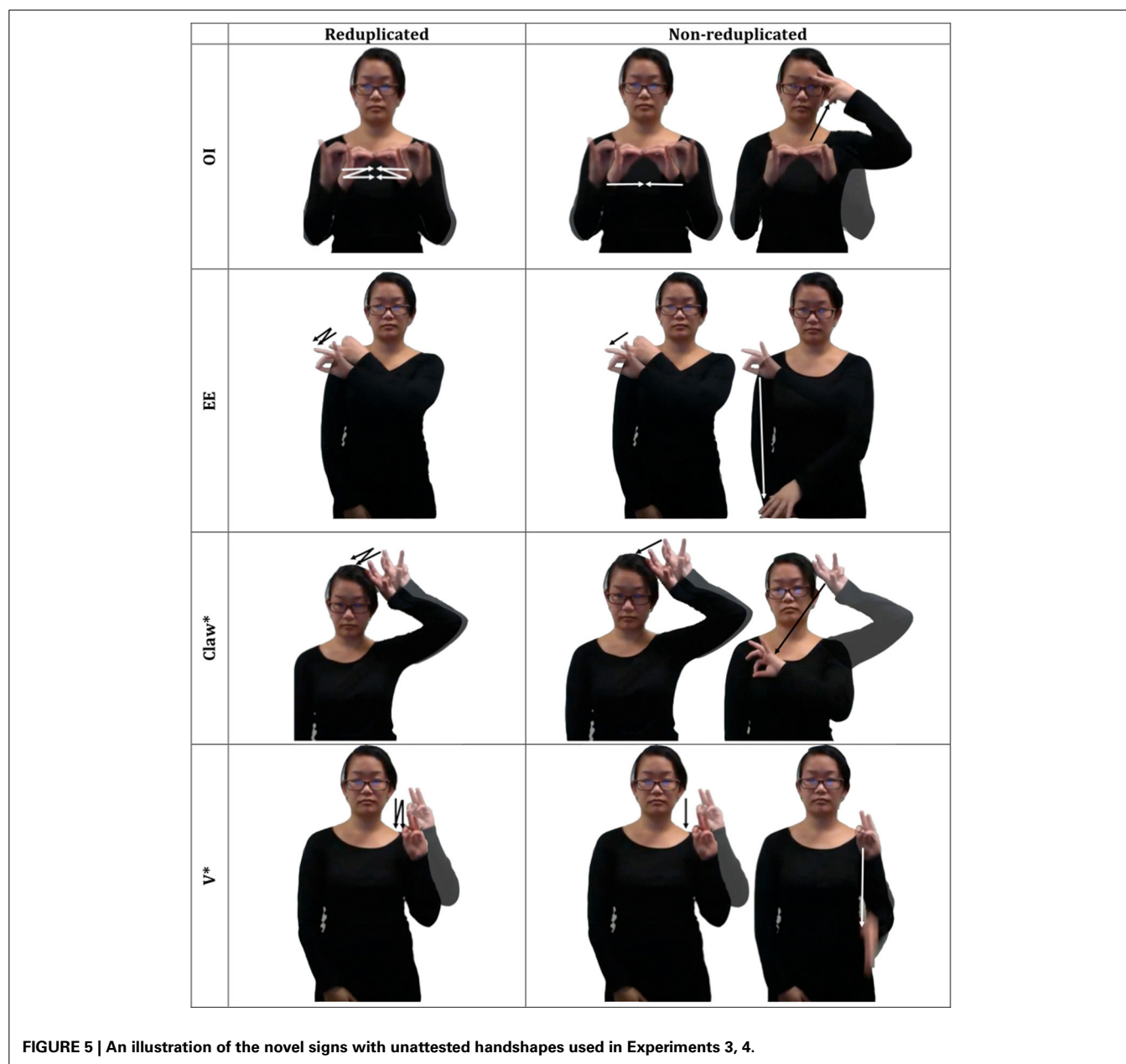
novel signs) ANOVAs. As in Experiment 2, response time was inspected to eliminate outliers (correct responses slower than 3000 ms or faster than 250 ms, less than 1% of the total correct responses).

An inspection of the means (see **Figure 7**) suggests that the unattested handshapes in Experiment 4 elicited faster and more accurate responses. While these savings were evident irrespective of lexicality, their magnitude was stronger for novel signs relative to ASL signs. Accordingly, the ANOVAs yielded reliable effects of attestation [In errors: $F1_{(1, 10)} = 37.34$, MSE $= 0.003$, $p < 0.0002$; $F2_{(1, 30)} = 5.87$, MSE $= 0.076$, $p < 0.03$; In response time: $F1_{(1, 10)} = 42.17$, MSE $= 21,632$, $p < 0.00001$; $F2_{(1, 30)} = 31.67$, MSE $= 14,538$, $p < 0.00001$] and lexicality [In errors: $F1_{(1, 10)} = 4.45$, MSE $= 0.003$, $p < 0.07$;

$F2_{(1, 30)} = 21.42$, MSE $= 0.033$, $p < 0.0001$; In response time: $F1_{(1, 10)} = 47.31$, MSE $= 7554$, $p < 0.00001$; $F2_{(1, 30)} = 343.92$, MSE $= 3841$, $p < 0.0001$]. The interaction was significant in the analyses of response time [$F1_{(1, 10)} = 34.80$, MSE $= 1576$, $p < 0.0002$; $F2_{(1, 30)} = 17.65$, MSE $= 3481$, $p < 0.0003$], and marginally significant in errors [$F1_{(1, 10)} = 15.1$, MSE $= 0.002$, $p < 0.004$; $F2_{(1, 30)} = 2.53$, MSE $= 0.033$, $p < 0.13$].

Tukey HSD tests showed that responses to ASL signs were significantly faster in the presence of novel signs with unattested handshapes compared to ones with attested handshapes ($p < 0.001$, by participants and items). Likewise, novel signs with unattested handshapes elicited faster and more accurate responses relative to those with attested handshapes ($p < 0.001$, by participants and items).

FIGURE 6 | Rating preference for reduplicated signs with unattested handshapes in Experiment 3.
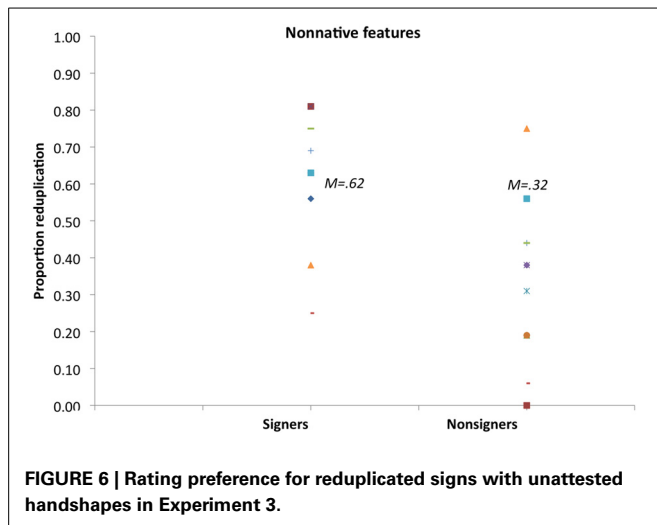
Having established that participants did notice the presence of unattested handshapes, we can next ask whether they represented their reduplicative structure. To this end, we now turn to examine the effect of reduplication on responses to ASL signs and novel signs in Experiment 4.

Are signers sensitive to the reduplication of unattested hand-shapes?

Errors. An inspection of the means (see **Figure 8**) suggests that reduplication produced different effects for existing signs and novel signs. The 2 reduplication × 2 lexicality ANOVAs on the proportion of errors (arcsine transformed) only produced a marginally significant interaction [$F1_{(1, 11)} = 4.89$, MSE $= 0.035$, $p < 0.05$; $F2_{(1, 30)} = 1.67$, MSE $= 0.058$, $p < 0.21$][9].

A simple main effect analysis demonstrated that novel reduplicated signs elicited a significant increase in errors relative to nonreduplicated controls [$t1_{(11)} = 2.73$, $p < 0.02$; $t2_{(15)} = 1.84$, $p < 0.05$ one-tailed]. In contrast, for attested ASL signs, reduplication resulted in a nonsignificant decrease in errors (both $t < 1$).

Response time. An inspection of the means (see **Figure 8**) suggests that reduplication facilitated response time for both signs and nonsigns, although this effect appears more pronounced for attested ASL signs.

The 2 lexicality × 2 reduplication ANOVAs yielded reliable effects of lexicality [$F1_{(1, 11)} = 12.82$, MSE $= 10,534$, $p < 0.005$; $F2_{(1, 30)} = 9.84$, MSE $= 17,765$, $p < 0.004$], reduplication [$F1_{(1, 11)} = 55.12$, MSE $= 3374$, $p < 0.0001$; $F2_{(1, 30)} = 23.06$, MSE $= 10,238$, $p < 0.0005$] and their interaction [$F1_{(1, 11)} = 16.98$, MSE $= 2557$, $p < 0.002$; $F2_{(1, 30)} = 6.05$, MSE $= 10,238$, $p < 0.02$]. The simple main effect of reduplication was significant for both signs [$t1_{(11)} = 9.19$, $p < 0.0001$; $t2_{(15)} = 4.65$, $p < 0.0004$] and novel signs [$t1_{(11)} = 2.66$, $p < 0.03$; $t2_{(15)} = 1.87$, $p < 0.05$, one-tailed].

---

[9]The interaction was likewise reliable in the logit analysis ($\beta = -0.627$, SE $= 0.319$, $Z = -1.97$, $p < 0.05$).

FIGURE 7 | The effect of handshape attestation on lexical decision across experiments. Note: Error bars are 95% confidence intervals for the difference between the means.
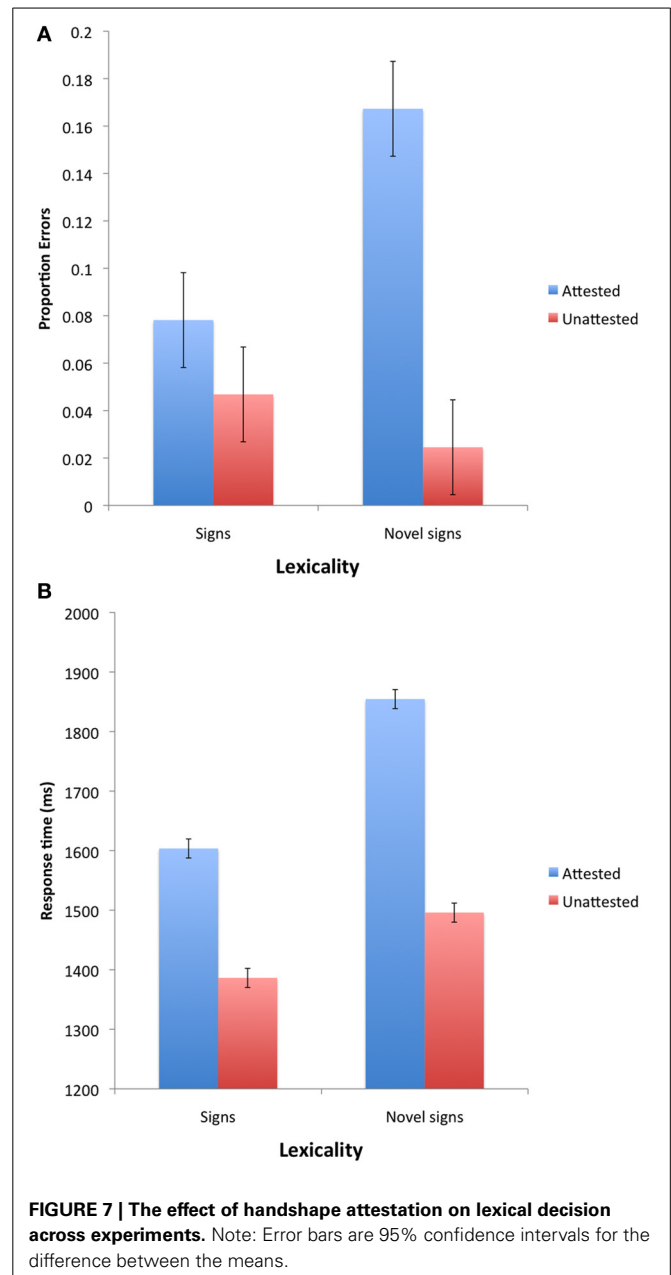
Discussion

The main finding of Experiment 4 is that signers are sensitive to the structure of novel signs with unattested ASL handshapes. First, participants had registered the presence of unattested hand-shapes, as their lexical decision responses in this experiment (i.e., in the presence of unattested handshapes) were reliably faster and more accurate relative to Experiment 2 (where all stimuli had handshapes that are native to ASL)[10]. Crucially, participants

---

[10]The ease of discrimination in Experiment 4 is unlikely to reflect a simple practice effect (due to its administration after Experiment 2) as a median split analysis of response accuracy in Experiment 2 and 4 according to trial order (first vs. second half) found no reliable effects of block order ($t < 1$).

FIGURE 8 | Lexical decision results for ASL signs and novel signs with handshapes in Experiment 4. Note: Error bars are 95% confidence intervals for the difference between the means.

An inspection of the materials indeed showed that the duration of reduplicated stimuli were overall shorter than nonreduplicated stimuli for both ASL signs ($M = 2024$ ms, $M = 2054$ ms; for reduplicated and nonreduplicated signs, respectively) and novel signs ($M = 2168$ ms, $M = 2201$ ms; for reduplicated and nonreduplicated signs, respectively). While this difference may well reflect a systematic effect of reduplication on sign production, its presence confounds the effect of reduplication on perception.

To address this limitation, we assessed the effect of reduplication in a stepwise linear regression analysis, conducted separately for ASL signs and novel signs. Stimulus duration was forced into the model in the first step; reduplication was entered last. Results showed that, for existing ASL signs, the effect of reduplication remained highly significant, even after controlling for the effect of stimulus duration [$R^2_{change} = 0.318$, $F2_{(1, 29)} = 20.64$, $p < 0.0001$] In contrast, once stimulus duration was controlled, the effect of reduplication on novel signs was no longer significant [$R^2_{change} = 0.049$, $F2_{(1, 29)} = 1.93$, $p < 0.18$, n.s.][11].

Together, the results establish that reduplicated signs are identified as more sign-like. Existing ASL signs that exhibit reduplication are identified more rapidly than nonreduplicated controls. Crucially, reduplication exerts the opposite effect for novel signs. Once stimulus duration was controlled, reduplication did not affect response time, but it reliably elevated errors to novel reduplicated signs. These findings demonstrate that signers extracted reduplication of novel features that they have never encountered before. This conclusion is consistent with the possibility that ASL signers encode abstract algebraic rules.

GENERAL DISCUSSION

Spoken languages include productive principles that allow speakers to extend their linguistic knowledge to novel instances (Chomsky, 1957). Across-the-board generalizations are significant because they are the hallmark of abstract algebraic rules (Fodor and Pylyshyn, 1988; Pinker and Prince, 1988; Marcus, 2001). Here, we asked whether such rules might also form part of the computational machinery of sign language. To this end, we examined whether signers can likewise extend their linguistic knowledge broadly.

As a case study, we examined signers' capacity to extend a reduplication rule—a rule that *inter alia* forms disyllabic nouns by reduplicating their monosyllabic verbal bases (X→XX). In four experiments, we asked whether signers extend reduplication to novel signs. Experiments 1–2 examined novel signs that reduplicate native ASL syllables; in Experiments 3–4, we probed for

were sensitive to the reduplicative structure of these stimuli. Novel reduplicated signs produced a higher error rate compared to nonreduplicated controls. In contrast, reduplicated ASL signs elicited faster responses.

The selectivity of the effect of reduplication to the lexicality of the stimulus—whether it is an ASL sign or a novel sign—would appear to suggest that reduplicated signs are generally identified as more sign-like. Consequently, reduplication renders novel signs harder to classify as such. This conclusion, however, is countered by the finding that the response time saving associated with reduplication extended even for novel signs. Thus, for novel signs, reduplication elevated error rates, but sped up response time.

These conflicting effects of reduplication on response time and accuracy are amenable to two distinct explanations. One possibility is that reduplication incurs genuine savings in the processing of novel signs—perhaps because the redundancy facilitates their encoding by the visual system. Alternatively, the effect of reduplication could emanate from uncontrolled variations in the duration of these stimuli.

---

[11]Another alternative explanation attributes the effect of reduplication to uncontrolled variation in movement repetitions. Since some of our nonreduplicated (XY) controls did not share the same movement type in their X and Y syllables, the co-occurrence of two identical movement types could have rendered novel reduplicated signs more sign-like. Most (11/16) item pairs, however, did share the same movement type. Moreover, a comparison of item pairs that shared the same movement type with those that did not (via a 2 movement × 2 reduplication ANOVA) found no effect of movement repetition on response accuracy (all $F < 1$). Accordingly, the effect of reduplication is unlikely due to the type of movement alone.

the reduplication of syllables whose handshape features are unattested in ASL. Given that reduplicated disyllables are favored in ASL (i.e., they are more frequent and possibly unmarked relative to nonreduplicated disyllables), we expected the reduplication rule to elicit a preference for novel reduplicated signs. This prediction was borne out in each of our four experiments. Experiments 1 and 3 showed that novel reduplicated signs are preferred to their nonreduplicative counterparts, and this preference obtained irrespective of whether the reduplicative feature is attested in ASL (in Experiment 1) or unattested (Experiment 3). Experiments 2 and 4 demonstrated that signers encode reduplication on-line, in lexical decision. In both experiments, novel signs with reduplicated features were more difficult to identify than their nonreduplicated counterparts, whereas reduplicated signs were identified more readily.

It is unlikely that the preference for reduplicated signs reflects a generic perceptual advantage. In fact, reduplicative signs were systematically dispreferred by nonsigners (in Experiments 1 and 3), and they were harder for signers to process (for novel signs, in Experiments 2 and 4).

The preference for reduplicated syllables is likewise inexplicable by their feature similarity (i.e., the fact that the XX syllables shared all their features, whereas XY syllables only shared some of those features). Our survey of nonreduplicative disyllables in the ASL lexicon reveals that partly similar signs—those in which the two syllables share location—are systematically *underrepresented* relative to dissimilar signs (i.e., those in which the location feature is not shared; for details, see footnote 12)[12]. Thus, acceptability (estimated by lexical frequency) is not a linear function of similarity (i.e., feature overlap): full identity is preferred, but partial similarity is systematically avoided—a result also found in spoken languages (e.g., Berent and Shimron, 2003; Berent et al., 2004). This conclusion counters the possibility that the preference for reduplicated signs (most critically, ones with unattested handshapes) is only due to the partial similarity among some of their native features. Further evidence against this possibility is presented by responses to the nonreduplicative disyllables in our experiments. Had the preference for XX signs been solely due to the (partial) overlap among their native features, then feature overlap should have predicted the acceptability of nonreduplicative XY signs—novel XY with greater feature overlap should have appeared more sign-like, hence, harder to identify as novel signs.

However, our results yield no correlation between the acceptability of novel XY signs (across Experiments 2 and 4) and their feature similarity [$r_{(30)} = 0.08$, for both accuracy and response time]. Given that partial similarity appears to be dispreferred (as judged by its underrepresentation in the lexicon), the preference for XX signs must be specifically due to the full identity of their syllables, including their unattested handshape.

Another similarity-based explanation attributes the preference for XX signs to the statistical properties of the ASL lexicon. But this explanation is also inconsistent with the available evidence. Recall that in Experiments 3–4, XX signs were favored to XY controls despite having two unattested handshapes (compared to only one unattested handshape in the XY controls). Thus, the preference for reduplicated signs is irreducible to their feature similarity to ASL signs. It is also unlikely that novel XX signs had larger neighborhoods than XY signs. By definition, XX signs with two unattested handshapes have no neighbors at all, as a neighbor differs from the target on a single parameter (Baus et al., 2008; Carreiras et al., 2008). Likewise, the neighborhoods of our attested signs were extremely sparse, as only two of our items had a neighbor (one reduplicated, with a single neighbor, and one nonreduplicated, with two neighbors). These observations offer no support for the lexical similarity account. Given that the preference for XX syllables is inexplicable by either the feature similarity among their two syllables or their statistical similarity to the ASL lexicon, the most likely explanation for our results is that the preference for XX signs reflects their reduplication.

Our findings show for the first time that signers' knowledge of their native language supports systematic generalizations that extend across the board—even to features that they have never encountered before. Algebraic rules provide a natural computational explanation for these findings. Because such rules operate on variables that stand for entire equivalence classes (e.g., any syllables), algebraic rules apply broadly, irrespective of the familiarity with novel items and their similarity to familiar stimuli.

Not only are these results consistent with the encoding of algebraic rules, they are also inconsistent with a nonalgebraic alternative. Past computational simulations, attempting to capture reduplication rules using nonalgebraic mechanisms (i.e., mechanisms that lack the capacity to operate over variables)—either connectionist networks (Marcus, 1998, 2001), or a state of the art inductive learner (Berent et al., 2012b)—have failed to adequately capture human generalizations. As in the present experiment, these simulations examined generalization of an identity function to test items including a single unattested feature. Results showed that, absent operations on variables, these models failed to generalize to such items. While the capacity of such models to account for the present data remains to be seen, the close parallels with previous test cases from spoken language suggest that their success for reduplicated signs is unlikely. Accordingly, signers' capacity to extend reduplication across the board suggests that their linguistic knowledge of reduplication relies on algebraic rules.

The conclusion that the ASL grammar encodes algebraic rules does not speak to the precise nature of rules available to participants. While our materials were modeled after the morphological

---

[12]To determine whether partial feature similarity is preferred, we extracted from an on-line ASL dictionary (ASLpro.com) all disyllabic signs whose two syllables are nonreduplicative—a total of 366 signs. To isolate the effect of feature overlap along a single parameter—location—we further limited the search to nonreduplicative signs whose syllables do not share a handshape—a total of 188 signs. We next coded each such sign for the location of its two syllables along ten different location categories (mouth, neutral, head, contact with non-dominant hand, chest, arm, ear, face, chin, torso), and indicated whether or not the two syllables share the same location. Of the 188 signs surveyed, only 33 signs (i.e., 0.175) shared location—a proportion that is unexpected by the chance level of 0.5 ($p < 0.0001$ by a binomial test). This finding demonstrates that, in the absence of full identity, partial feature similarity is actively avoided in the ASL lexicon. This finding is inconsistent with the possibility that the preference for reduplicated signs in our experiments is due to the partial feature overlap among the two syllables.

rule that obtains nouns from verb reduplication, these results cannot determine whether signers effectively represented the novel reduplicative signs as nouns. We also note that our evidence for rules does not negate the possibility that some aspects of linguistic knowledge are associative, or even iconic (Ormel et al., 2009; Thompson et al., 2009, 2010, 2012). While these alternative representations and computational mechanisms might be ultimately necessary to offer a full account of the language system, our present results suggest that they are not sufficient. At its core, signers' phonological knowledge includes productive algebraic rules, akin to the ones previously documented in spoken language phonology. These results suggest that the computational architecture of the phonological mind is at least partly amodal (Berent, 2013a,b).

AUTHOR NOTES

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://www.frontiersin.org/journal/10.3389/fpsyg.2014.00560/abstract

REFERENCES

Baker, S. A., Idsardi, W. J., Golinkoff, R. M., and Petitto, I.-A. (2005). The perception of handshapes in American sign language. *Mem. Cogn.* 33, 887–904. doi: 10.3758/BF03193083

Baker, S. A., Michnick Golinkoff, R., and Petitto, L.-A. (2006). New insights into old puzzles from infants' categorical discrimination of soundless phonetic units. *Lang. Learn. Dev.* 2, 147–162. doi: 10.1207/s15473341lld0203_1

Baus, C., Gutiérrez-Sigut, E., Quer, J., and Carreiras, M. (2008). Lexical access in Catalan Signed Language (LSC) production. *Cognition* 108, 856–865. doi: 10.1016/j.cognition.2008.05.012

Berent, I. (2013a). *The Phonological Mind.* Cambridge: Cambridge University Press.

Berent, I. (2013b). The phonological mind. *Trends Cogn. Sci.* 17, 319–327. doi: 10.1016/j.tics.2013.05.004

Berent, I., Balaban, E., and Vaknin-Nusbaum, V. (2011). How linguistic chickens help spot spoken-eggs: phonological constraints on speech identification. *Front. Psychol.* 2:182. doi: 10.3389/fpsyg.2011.00182

Berent, I., Bibi, U., and Tzelgov, J. (2006). The autonomous computation of linguistic structure in reading: evidence from the Stroop task. *Ment. Lex.* 1, 201–230. doi: 10.1075/ml.1.2.03ber

Berent, I., Dupuis, A., and Brentari, D. (2013). Amodal aspects of linguistic design. *PLoS ONE* 8:e60617. doi: 10.1371/journal.pone.0060617

Berent, I., Everett, D. L., and Shimron, J. (2001a). Do phonological representations specify variables? Evidence from the obligatory contour principle. *Cogn. Psychol.* 42, 1–60. doi: 10.1006/cogp.2000.0742

Berent, I., Marcus, G. F., Shimron, J., and Gafos, A. I. (2002). The scope of linguistic generalizations: evidence from Hebrew word formation. *Cognition* 83, 113–139. doi: 10.1016/S0010-0277(01)00167-6

Berent, I., and Shimron, J. (1997). The representation of Hebrew words: evidence from the obligatory contour principle. *Cognition* 64, 39–72. doi: 10.1016/S0010-0277(97)00016-4

Berent, I., and Shimron, J. (2003). Co-occurrence restrictions on identical consonants in the Hebrew lexicon: are they due to similarity? *J. Linguist.* 39, 31–55. doi: 10.1017/S0022226702001949

Berent, I., Shimron, J., and Vaknin, V. (2001b). Phonological constraints on reading: evidence from the Obligatory Contour Principle. *J. Mem. Lang.* 44, 644–665. doi: 10.1006/jmla.2000.2760

Berent, I., Vaknin, V., and Shimron, J. (2004). Does a theory of language need a grammar? Evidence from hebrew root structure. *Brain Lang.* 90, 170–182. doi: 10.1016/S0093-934X(03)00430-9

Berent, I., Vaknin-Nusbaum, V., Balaban, E., and Galaburda, A. M. (2012a). Dyslexia impairs speech recognition but can spare phonological competence. *PLoS ONE* 7:e44875. doi: 10.1371/journal.pone.0044875

Berent, I., Wilson, C., Marcus, G., and Bemis, D. (2012b). On the role of variables in phonology: remarks on Hayes and Wilson. *Linguist. Inq.* 43, 97–119. doi: 10.1162/LING_a_00075

Best, C. T., Mathur, G., Miranda, K. A., and Lillo-Martin, D. (2010). Effects of sign language experience on categorical perception of dynamic ASL pseudosigns. *Attent. Percept. Psychophys.* 72, 747–762. doi: 10.3758/APP.72.3.747

Bochner, J. H., Christie, K., Hauser, P. C., and Searls, J. M. (2011). When is a difference really different? Learners' discrimination of linguistic contrasts in American sign language. *Lang. Learn.* 61, 1302–1327. doi: 10.1111/j.1467-9922.2011.00671.x

Bosworth, R. G., and Emmorey, K. (2010). Effects of iconicity and semantic relatedness on lexical access in american sign language. *J. Exp. Psychol. Learn. Mem. Cogn.* 36, 1573–1581. doi: 10.1037/a0020934

Brentari, D. (1993). Establishing a sonority hierarchy in American Sign Language: the use of simultaneous structure in phonology. *Phonology* 10, 281–306. doi: 10.1017/S0952675700000063

Brentari, D. (1994). "Prosodic constraints in American Sign Language," in *Sign Language Research*, eds H. Bos and T. Schermer (Hamburg: Signum Press), 39–51.

Brentari, D. (1998). *A Prosodic Model of Sign Language Phonology.* Cambridge, MA: MIT Press.

Brentari, D. (2006). "Effects of language modality on word segmentation: an experimental study of phonological factors in a sign language," in *Papers in Laboratory Phonology VIII*, eds L. Goldstein, D. Whalen, and C. Best (Berlin: Mouton de Gruyter), 155–164.

Brentari, D. (2007). "Sign language phonology: issues of iconicity and universality," in *Verbal and Signed Languages*, eds E. Pizzuto and R. Simone (Berlin: Mouton de Gruyter), 59–80.

Brentari, D. (2011). "Sign language phonology," in *Handbook of Phonological Theory*, eds J. Goldsmith, J. Riggle, and A. Yu (New York; Oxford: Blackwells), 691–721.

Brentari, D., González, C., Seidl, A., and Wilbur, R. (2011). Sensitivity to visual prosodic cues in signers and nonsigners. *Lang. Speech* 54, 49–72. doi: 10.1177/0023830910388011

Bybee, J. L. (2008). "Linguistic universals and language change," in *Linguistic Universals and Language Change*, ed J. Good (Oxford; New York: Oxford University Press), 108–121.

Bybee, J., and McClelland, J. L. (2005). Alternatives to the combinatorial paradigm of linguistic theory based on domain general principles of human cognition. *Linguist. Rev.* 22, 381–410. doi: 10.1515/tlir.2005.22.2-4.381

Carreiras, M., Gutiérrez-Sigut, E., Baquero, S., and Corina, D. (2008). Lexical processing in Spanish sign language (LSE). *J. Mem. Lang.* 58, 100–122. doi: 10.1016/j.jml.2007.05.004

Chomsky, N. (1957). *Syntactic Structures.* Gravenhage: Mouton.

Chomsky, N. (1980). *Rules and Representations.* New York, NY: Columbia University Press.

Chomsky, N., and Halle, M. (1968). *The Sound Pattern of English.* New York, NY: Harper and Row.

Corina, D. P. (1990). "Reassessing the role of sonority in syllable structure: evidence from visual gestural language," in *Papers From the 26th Annual Regional Meeting of the Chicago Linguistic Society*, Vol. 2, eds M. Ziolkowski, M. Noske, and K. Deaton (Chicago, IL: University of Chicago), 33–43.

Corina, D. P., and Hildebrandt, U. C. (2002). "Psycholinguistic investigations of phonological structure in ASL," in *Modality and Structure in Signed and Spoken Languages* (New York, NY: Cambridge University Press), 88–111.

Corina, D., and Sandler, W. (1993). On the nature of phonological structure in sign language. *Phonology* 10, 165–207. doi: 10.1017/S0952675700000038

de Lacy, P. (2006). *Markedness: Reduction and Preservation in Phonology.* Cambridge; NY: Cambridge University Press. doi: 10.1017/CBO9780511486388

Eccarius, P., and Brentari, D. (2010). A formal analysis of phonological contrast and iconicity in sign language handshapes. *Sign Lang. Linguist.* 13, 156–181. doi: 10.1075/sll.13.2.02ecc

Elman, J. (1993). Learning and development in neural networks: the importance of starting small. *Cognition* 48, 71–99. doi: 10.1016/0010-0277(93)90058-4

Elman, J. L. (2005). Connectionist models of cognitive development: where next? *Trends Cogn. Sci.* 9, 111–117. doi: 10.1016/j.tics.2005.01.005

Elman, J. L., Bates, E. A., Johnson, M. H., Karmiloff-Smith, A., Parisi, D., and Plunkett, K. (1996). *Rethinking Innateness: a Connectionist Perspective on Development*. Cambridge: MIT press.

Emmorey, K., and Corina, D. (1990). Lexical recognition in sign language: effects of phonetic structure and morphology. *Percept. Mot. Skills* 71, 1227–1252. doi: 10.2466/pms.1990.71.3f.1227

Emmorey, K., Grabowski, T., McCullough, S., Damasio, H., Ponto, L., Hichwa, R., et al. (2004). Motor-iconicity of sign language does not alter the neural systems underlying tool and action naming. *Brain Lang.* 89, 27–37. doi: 10.1016/S0093-934X(03)00309-2

Emmorey, K., McCullough, S., and Brentari, D. (2003). Categorical perception in American Sign Language. *Lang. Cogn. Process.* 18, 21–45. doi: 10.1080/01690960143000416

Endress, A. D., Scholl, B. J., and Mehler, J. (2005). The role of salience in the extraction of algebraic rules. *J. Exp. Psychol. Gen.* 134, 406–419. doi: 10.1037/0096-3445.134.3.406

Fodor, J., and Pylyshyn, Z. (1988). Connectionism and cognitive architecture: a critical analysis. *Cognition* 28, 3–71. doi: 10.1016/0010-0277(88)90031-5

Frampton, J. (2009). *Distributed Reduplication*. Cambridge, MA: MIT Press. doi: 10.7551/mitpress/9780262013260.001.0001

Gafos, A. I. (1999). *The Articulatory Basis of Locality in Phonology*. New York, NY: Garland publishers.

Gervain, J., Berent, I., and Werker, J. (2012). Binding at birth: newborns detect identity relations and sequential position in speech. *J. Cogn. Neurosci.* 24, 564–574. doi: 10.1162/jocn_a_00157

Gervain, J., Macagno, F., Cogoi, S., Peña, M., and Mehler, J. (2008). The neonate brain detects speech structure. *Proc. Natl. Acad. Sci. U.S.A.* 105, 14222–14227. doi: 10.1073/pnas.0806530105

Greenberg, J. H. (1950). The patterning of morphemes in Semitic. *Word* 6, 162–181.

Gutiérrez, E., Müller, O., Baus, C., and Carreiras, M. (2012). Electrophysiological evidence for phonological priming in Spanish Sign Language lexical access. *Neuropsychologia* 50, 1335–1346. doi: 10.1016/j.neuropsychologia.2012.02.018

Haskell, T. R., MacDonald, M. C., and Seidenberg, M. S. (2003). Language learning and innateness: some implications of compounds research. *Cogn. Psychol.* 4, 119–163. doi: 10.1016/S0010-0285(03)00007-0

Hayes, B., and Wilson, C. (2008). A maximum entropy model of phonotactics and phonotactic learning. *Linguist. Inq.* 39, 379–440. doi: 10.1162/ling.2008.39.3.379

Hildebrandt, U., and Corina, D. (2002). Phonological similarity in american sign language. *Lang. Cogn. Process.* 17, 593–612. doi: 10.1080/01690960143000371

Jantunen, T., and Takkinen, R. (2010). "Syllable structure in sign language phonology," in *Sign Languages*, ed D. Brentari (Cambridge: Cambridge University Press), 312–331.

Klima, E. S., and Bellugi, U. (1979). *The Signs of Language. Cambridge*, MA: Harvard University Press.

Lane, H., Boyes-Braem, P., and Bellugi, U. (1976). Preliminaries to a distinctive feature analysis of handshapes in American sign language. *Cogn. Psychol.* 8, 263–289. doi: 10.1016/0010-0285(76)90027-X

Leben, W. (1973). *Suprasegmental Phonology*. Cambridge, MA: MIT press.

Mahon, B. Z., and Caramazza, A. (2008). A critical look at the embodied cognition hypothesis and a new proposal for grounding conceptual content. *J. Physiol. Paris* 102, 59–70. doi: 10.1016/j.jphysparis.2008.03.004

Marcus, G. F. (1998). Rethinking eliminative connectionism. *Cogn. Psychol.* 37, 243–282. doi: 10.1006/cogp.1998.0694

Marcus, G. (2001). *The Algebraic Mind: Integrating Connectionism and Cognitive Science*. Cambridge: MIT press.

Marcus, G. F., Fernandes, K. J., and Johnson, S. P. (2007). Infant rule learning facilitated by speech. *Psychol. Sci.* 18, 387–391. doi: 10.1111/j.1467-9280.2007.01910.x

Marcus, G. F., Vijayan, S., Bandi Rao, S., and Vishton, P. M. (1999). Rule learning by seven-month-old infants. *Science* 283, 77–80. doi: 10.1126/science.283.5398.77

Marentette, P. F., and Mayberry, R. I. (2000). "Principles for an emerging phonological system: a case study of early ASL acquisition," in *Language Acquisition by Eye* (Mahwah, NJ: Lawrence Erlbaum Associates Publishers), 71–90.

Marshall, C. R., Denmark, T., and Morgan, G. (2006). Investigating the underlying causes of SLI: a non-sign repetition test in British Sign Language. *Adv. Speech Lang. Pathol.* 8, 347–355. doi: 10.1080/14417040600970630

McCarthy, J. J. (1986). OCP effects: gemination and antigemination. *Linguist. Inq.* 17, 207–263.

McCarthy, J. J. (1989). Linear order in phonological representation. *Linguist. Inq.* 20, 71–99.

McCarthy, J. J., and Prince, A. (1995). "Prosodic morphology," in *Phonological Theory*, ed J. A. Goldsmith (Oxford: Basil Blackwell), 318–366.

McClelland, J. L. (2009). "Phonology and perception: a cognitive scientist's perspective," in *Phonology in Perception*, eds P. Boersma and S. Hamann (Berlin: Mouton De Gruyter), 293–314.

McClelland, J. L., Botvinick, M. M., Noelle, D. C., Plaut, D. C., Rogers, T. T., Seidenberg, M. S., et al. (2010). Letting structure emerge: connectionist and dynamical systems approaches to cognition. *Trends Cogn. Sci.* 14, 348–356. doi: 10.1016/j.tics.2010.06.002

McClelland, J. L., and Patterson, K. (2002). Rules or connections in past-tense inflections: what does the evidence rule out? *Trends Cogn. Sci.* 6, 465–472. doi: 10.1016/S1364-6613(02)01993-9

McClelland, J., and Plaut, D. (1999). Does generalization in infant learning implicate abstract algebra-like rules? *Trends Cogn. Sci.* 3, 166–168. doi: 10.1016/S1364-6613(99)01320-0

Meier, R. P. (2000). "Shared motoric factors in the acquisition of sign and speech," in *The Signs of Language Revisited: An Anthology to Honor Ursula Bellugi and Edward Klima*, eds K. Emmorey and H. Lane (Mahwah, NJ: Lawrence Erlbaum Associates Publishers), 333–356.

Meier, R. P., Mauk, C. E., Cheek, A., and Moreland, C. J. (2008). The form of children's early signs: Iconic or motoric determinants. *Lang. Learn. Dev.* 4, 63–98. doi: 10.1080/15475440701377618

Morgan, G. (2006). Children are just lingual: the development of phonology in British Sign Language (BSL). *Lingua* 116, 1507–1523. doi: 10.1016/j.lingua.2005.07.010

Morgan, G., Barrett-Jones, S., and Stoneham, H. (2007). The first signs of language: phonological development in British Sign Language. *Appl. Psycholinguist.* 28, 3–22. doi: 10.1017/S0142716407070014

Newport, E. (1982). "Task specificity in language learning? Evidence from speech perception and American Sign Language," in *Language Acquisition: the State of the Art*, eds E. Wanner and L. Gleitman (Cambridge: Cambridge University Press), 450–486.

Orfanidou, E., Adam, R., McQueen, J. M., and Morgan, G. (2009). Making sense of nonsense in British Sign Language (BSL): the contribution of different phonological parameters to sign recognition. *Mem. Cogn.* 37, 302–315. doi: 10.3758/MC.37.3.302

Orfanidou, E., Adam, R., Morgan, G., and McQueen, J. M. (2010). Recognition of signed and spoken language: different sensory inputs, the same segmentation procedure. *J. Mem. Lang.* 62, 272–283. doi: 10.1016/j.jml.2009.12.001

Ormel, E., Knoors, H., Hermans, D., and Verhoeven, L. (2009). The role of sign phonology and iconicity during sign processing: the case of deaf children. *J. Deaf Stud. Deaf Educ.* 14, 436–448. doi: 10.1093/deafed/enp021

Palmer, S. B., Fais, L., Golinkoff, R. M., and Werker, J. F. (2012). Perceptual narrowing of linguistic sign occurs in the 1st year of life. *Child Dev.* 83, 543–553. doi: 10.1111/j.1467-8624.2011.01715.x

Perlmutter, D. M. (1992). Sonority and syllable structure in American Sign Language. *Linguist. Inq.* 23, 407–442.

Pinker, S. (1994). *The Language Instinct*. New York, NY: Morrow.

Pinker, S. (1999). *Words and Rules: the Ingredients of Language*. New York, NY: Basic Books.

Pinker, S., and Prince, A. (1988). On language and connectionism: analysis of a parallel distributed processing model of language acquisition. *Cognition* 28, 73–193. doi: 10.1016/0010-0277(88)90032-7

Poizner, H., Bellugi, U., and Lutes-Driscoll, V. (1981). Perception of American Sign Language in dynamic point-light displays. *J. Exp. Psychol. Hum. Percept. Perform.* 7, 430–440. doi: 10.1037/0096-1523.7.2.430

Prince, A., and Smolensky, P. (1993/2004). *Optimality Theory: Constraint Interaction in Generative Grammar*. Malden, MA: Blackwell Pub.

Rabagliati, H., Senghas, A., Johnson, S., and Marcus, G. F. (2012). Infant rule learning: advantage language, or advantage speech? *PLoS ONE* 7:e40517. doi: 10.1371/journal.pone.0040517

Ramscar, M., and Dye, M. (2011). Learning language from the input: why innate constraints can't explain noun compounding. *Cogn. Psychol.* 62, 1–40. doi: 10.1016/j.cogpsych.2010.10.001

Rumelhart, D. E., and McClelland, J. L. (1986). "On learning past tense of English verbs: implicit rules or parallel distributed processing?" in *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, Vol. 2, eds D. E. Rumelhart, J. L. McClelland, and T. P. R. Group (Cambridge, MA: MIT Press), 216–271.

Sandler, W. (2008). "The syllable in sign language: considering the other natural language modality," in *The Syllable in Speech Production*, eds B. L. Davis and K. Zajdó (New York, NY: Lawrence Erlbaum Associates), 379–408.

Sandler, W., and Lillo-Martin, D. C. (2006). *Sign Language and Linguistic Universals*. Cambridge: Cambridge University Press. doi: 10.1017/CBO9781139163910

Schmidtke, D., Conrad, M., and Jacobs, A. M. (2014). Phonological iconicity. *Front. Psychol.* 5, 1–6. doi: 10.3389/fpsyg.2014.00080

Seidenberg, M. S., and Jeffery, L. E. (1999). Do infants learn grammar with algebra or statistics. *Science* 284, 433. doi: 10.1126/science.284.5413.433f

Siedlecki, T., and Bonvillian, J. D. (1993). Phonological deletion revisited: Errors in young children's two-handed signs. *Sign Lang. Stud.* 80, 223–242. doi: 10.1353/sls.1993.0000

Stokoe, W. C. Jr. (1960). Sign language structure: an outline of the visual communication systems of the american deaf. *J. Deaf Stud. Deaf Educ.* 10, 3–37. doi: 10.1093/deafed/eni001

Supalla, T., and Newport, E. (1978). "How many seats in a chair? The derivation of nouns and verbs in American Sign Language," in *Understanding Language through Sign Language Research*, ed P. Siple (San Diego: Academic Press), 91–132.

Suzuki, K. (1998). *A Typological Investigation of Dissimilation*. Tucson, AZ: University of Arizona.

Thompson, R., Emmorey, K., and Gollan, T. H. (2005). 'Tip of the Fingers' experiences by deaf signers. *Psychol. Sci.* 16, 856–860. doi: 10.1111/j.1467-9280.2005.01626.x

Thompson, R. L., Vinson, D. P., and Vigliocco, G. (2009). The link between form and meaning in american sign language: lexical processing effects. *J. Exp. Psychol. Learn. Mem. Cogn.* 35, 550–557. doi: 10.1037/a0014547

Thompson, R. L., Vinson, D. P., and Vigliocco, G. (2010). The link between form and meaning in British sign language: effects of iconicity for phonological decisions. *J. Exp. Psychol. Learn. Mem. Cogn.* 36, 1017–1027. doi: 10.1037/a0019339

Thompson, R. L., Vinson, D. P., Woll, B., and Vigliocco, G. (2012). The road to language learning is iconic: evidence from British sign language. *Psychol. Sci.* 23, 1443–1448. doi: 10.1177/0956797612459763

Toro, J. M., Nespor, M., Mehler, J., and Bonatti, L. L. (2008). Finding words and rules in a speech stream: functional differences between vowels and consonants. *Psychol. Sci.* 19, 137–144. doi: 10.1111/j.1467-9280.2008.02059.x

Wilbur, R. (2012). "Sign Syllables," in *The Blackwell Companion to Phonology*, eds M. van Oostendorp, C. J. Ewen, E. Hume, and K. Rice (Blackwell Publishing), 1309–1334.

Wilbur, R. B. (2009). Productive reduplication in a fundamentally monosyllabic language. *Lang. Sci.* 31, 325–342. doi: 10.1016/j.langsci.2008.12.017

Reproducing American Sign Language sentences: cognitive scaffolding in working memory

Ted Supalla[1]*, Peter C. Hauser[2] and Daphne Bavelier[3,4]

[1] Sign Language Research Lab, Department of Neurology, Center for Brain Plasticity and Recovery, Georgetown University, Washington, DC, USA
[2] Department of American Sign Language and Interpreting Education, Deaf Studies Laboratory, National Technical Institute for the Deaf, Rochester Institute of Technology, Rochester, NY, USA
[3] Department of Psychology and Education Sciences, University of Geneva, Geneva, Switzerland
[4] Department of Brain and Cognitive Sciences, University of Rochester, Rochester, NY, USA

The American Sign Language Sentence Reproduction Test (ASL-SRT) requires the precise reproduction of a series of ASL sentences increasing in complexity and length. Error analyses of such tasks provides insight into working memory and scaffolding processes. Data was collected from three groups expected to differ in fluency: deaf children, deaf adults and hearing adults, all users of ASL. Quantitative (correct/incorrect recall) and qualitative error analyses were performed. Percent correct on the reproduction task supports its sensitivity to fluency as test performance clearly differed across the three groups studied. A linguistic analysis of errors further documented differing strategies and bias across groups. Subjects' recall projected the affordance and constraints of deep linguistic representations to differing degrees, with subjects resorting to alternate processing strategies when they failed to recall the sentence correctly. A qualitative error analysis allows us to capture generalizations about the relationship between *error pattern* and the cognitive scaffolding, which governs the sentence reproduction process. Highly fluent signers and less-fluent signers share common chokepoints on particular words in sentences. However, they diverge in heuristic strategy. Fluent signers, when they make an error, tend to preserve semantic details while altering morpho-syntactic domains. They produce syntactically correct sentences with equivalent meaning to the to-be-reproduced one, but these are not verbatim reproductions of the original sentence. In contrast, less-fluent signers tend to use a more linear strategy, preserving lexical status and word ordering while omitting local inflections, and occasionally resorting to visuo-motoric imitation. Thus, whereas fluent signers readily use top-down scaffolding in their working memory, less fluent signers fail to do so. Implications for current models of working memory across spoken and signed modalities are considered.

Keywords: American Sign Language, working memory, error analysis, verbatim recall, native signers, reproduction error, error type

INTRODUCTION

Current literature in psycholinguistics and cognitive science has deepened our understanding of the nature of short term memory (STM), but much work remains in the description and modeling of working memory, particularly for understanding the impact of modality on language processing. Working memory is generally considered to be a scaffolding for cognitive functions required to accomplish a task (Baddeley, 1995). However, debate goes on as to whether the layers of linguistic processing are modular or interactive (Fodor, 1983; Just and Carpenter, 1992) and whether STM is separable from working memory (Baddeley and Hitch, 2007). Research into the nature of STM in a signed language so far reveals that STM capacity for individual signs is not identical to the processes and capacity used in the recall of spoken words (Boutla et al., 2004; Bavelier et al., 2006). One may ask what this implies for the working memory of signers in processing sentences. One way to address this

question is to examine the way signers use working memory to process and retain ASL sentences. To pursue this line of research, we have examined ASL sentence reproduction, particularly the effect of bottleneck conditions on this task. We hypothesized that there are similar kinds of processes and constraints on working memory for processing ASL sentences and spoken sentences. Furthermore, we hypothesized that during cognitive and linguistic encoding and production, fluent signers make use of linguistic scaffolding and parsing options that are not available to signers with lower levels of competence. An error analysis of signers across a range of fluency levels supports these hypotheses, with generalizations from the data consistent with current models of language processing, supporting processes of grammatically constrained regeneration of conceptual content (Potter and Lombardi, 1990) and showing the effects of effortful "explicit" processing at the lexical level (Ronnberg et al., 2008) in less fluent signers.

Psycholinguistic investigations of both signed and spoken language have shown that performance on many types of working memory tasks interacts significantly with language fluency and age of acquisition (Newport and Meier, 1985; Newport, 1990). In a seminal study of ASL sentence shadowing and recall, Mayberry and Fischer (1989) showed error patterns which point to different types of processing by native and late learners of the language. Native signers' errors were predominantly lexical substitutions that had a semantic relationship to the target sign and were unrelated to its phonological form. In contrast, later-learning signers' errors were predominantly those with a formational or phonological relationship to the target, but not to its meaning[1]. While subjects from both groups made both types of errors, they produced these errors in strikingly different proportions. Morford (2002) has proposed that early language exposure enables automaticity of phonological processing, one factor which may account for the difference in relative proportion of error types. Thus, the Mayberry and Fischer (1989) error pattern data can be framed as an interaction between automatic linguistic processing and conceptual regeneration for sentence shadowing. In these terms, native signers make semantic errors consistent with automatic processing, storing the content in conceptual terms, whereas late learners and less fluent signers might be seen to make more superficial errors because of their more limited abilities to process through the phonology to achieve a deeper representation of the sentence. In both the Mayberry and Fischer study and the study reported on in this paper, the detailed examination of error patterns in response to controlled target stimuli by groups of signers differing in aspects of hearing status, age and signing background reveals processing strategies and details which can illuminate our understanding of models of working memory.

The methodology of our study on working memory differs from Mayberry and Fischer's (1989) shadowing task in several ways. First, we ran subjects from a variety of linguistic backgrounds and pooled subjects in specific population groups to examine error distribution. Second, whereas Mayberry and Fischer used short, easily remembered sentences, our stimuli ranged from short, easily recalled sentences to longer and/or more complex utterances. This difference in stimuli and task led to a greater number and variety of error types for the American Sign Language Sentence Reproduction Test (ASL-SRT) data. In particular, while the Mayberry and Fischer data analyses categorized only semantic and phonological substitution types, our data analysis also included syntactic substitutions such as changes in sign order and morphological alternations. This richness in turn

---

[1]A practical strategy in the absence of grammatical knowledge is revealed in the responses of a small pilot group of deaf college students who grew up without being exposed to sign language, learning ASL when entering college. Error analysis of these subjects' responses shows that they tend to use a strategy in which they attempt to copy the visual/motoric parameters of the signing stream. While this imitation may seem "correct" when the response resembles the stimulus, it often results in unintelligible signing. This sort of error has been noted in the literature. While Mayberry and Fischer (1989) did not include a non-signing "novice" group in their reported experiments, they do mention this sort of strategy among naive signers in a pilot group, calling it "hand waving" and differentiating it from the strategy and performance of even late-acquiring ASL signers.

has allowed us to provide a more detailed analysis of the various constraints and processes at play during working memory for linguistic materials, and contributes to the development of a processing model that highlights a number of new and key features in linguistic working memory.

The ASL-SRT assesses ASL language proficiency by asking subjects to repeat verbatim a 20-item series of ASL sentences. Over the course of the test, the sentences increase in length, number of propositions, and morphological complexity (Hauser et al., 2008). In the present study, we examined in more depth the data of selected adult and young Deaf subjects from this study, as well as incorporating data from an additional pool of hearing subjects who also took this test, but whose data was not included in Hauser et al. (2008). All subjects had deaf parents who used ASL in the home. While all subjects were exposed to American Sign Language in homes with Deaf parents, various other factors affect their ASL proficiency at the time of testing. Within the sign language community, there is variation in the age at which signers are first exposed to ASL and in the input they receive to the language. Moreover, there is also variation in fluency even among native signers. Fluency increases with age (young children as compared with older children and adults) and also varies according to the extent of immersion and use of the language. For example, Deaf native signers often differ from Hearing native signers in whether ASL is their dominant language and how much they use the language in daily life. As a result of such differences, signers may or may not demonstrate fluency and a high level of proficiency in the sign language to which they are exposed. Including both the Deaf of Deaf Adults (DDA) (those raised by Deaf parents) and Deaf of Deaf Youths (DDY) groups allows us to examine the effect of age upon reproduction skills, while including the Hearing of Deaf Adults (HDA) group allows us to contrast the performance of variously fluent hearing signers with the Deaf groups while keeping home language backgrounds constant. In this way, we avoided confounding hearing status with L2 language issues. Interestingly, we found little evidence of intrusion of English grammar in the pool of 75 signers. One DDY added English features occasionally. For example, he replaced the ASL sign HAVE-TO with an English-based sequence of signs glossed as HAVE TO. The variation in fluency among the hearing offspring of Deaf parents is similar to the range from highly fluent to semi-speaker in children from minority or immigrant ethnic group families where parents continue using their native language at home. Possible factors affecting their fluency are the number of deaf siblings, if any, and birth order of the HDA subject.

The reproduction accuracy of all signers was examined as well as the nature of their response. All subjects took a previous version of the test with 39-items, but we examined data only from responses to the 20 test items included in the current version of the test. The determining factors for eliminating test sentences were the measured redundancy of some test items that showed a similar level of complexity, and the potential for inconsistencies from dialectal variation for particular items. The analysis of this sample establishes the effectiveness of the reproduction task as a tool for measuring fluency in a sign language, showing that the test is indeed sensitive to the differences in linguistic structure of signing among signers of varying ages and fluency. Furthermore,

the analysis confirms that native signing raters can reliably differentiate the accuracy of reproduction across groups whom we would expect to differ in fluency with more technical linguistic assessments of grammatical structure.

In this article, we first provide the definitive description of the ASL-SRT. We then discuss the quantitative analyses performed on the three groups of native signing subjects who took the test. We also outline the method and results from qualitative analyses of the ASL-SRT responses from this same pool of 75 participants. The data reveal that signers' error types differ according to individuals' relative level of competence, as measured by their reproduction accuracy. The stimuli and task are sensitive to the subjects' differing levels of exposure and use of ASL, with performance analyses showing that signers varied in success in reproducing a target form, even in a short, single-clause sentences. Moreover, Deaf and Hearing signers who obtained higher reproduction accuracy scores made different sorts of errors than weak signers. Among less fluent signers, responses often include ungrammatical sign forms and/or sentences. Furthermore, errors are less predictable than those of more fluent signers as sentences increase in complexity. In more fluent signers, complex sentence targets trigger specific processing difficulties and predictable types of errors. The escalating demands of the reproduction task also result in clusters of various types of errors, which are useful for teasing out processing at the interface between the layers of processing and specific phrasal domains.

We developed the American Sign Language Sentence Reproduction Task (ASL-SRT) with the goal of establishing a standardized instrument that could be used across age and ability level to assess proficiency and fluency of signers. In the responses of subjects, we see differences in overall reproduction accuracy as a reflection of signers' various levels of sign language exposure, use and resultant fluency. In addition, we see differences in the types of errors made by signers of different fluency levels and backgrounds. From the perspective of a cognitive scientist, the precisely controlled data from the ASL-SRT provide an opportunity to examine the way signers use working memory to process and reproduce sentences.

The error patterns across variably fluent groups have implications for current models of working memory across spoken and signed modalities. That is, the conventional model of serial processing for non-sentence material can be replaced by a hierarchical model for working memory with parallel processing capabilities, a top-down scaffolding mechanism that assists sentence reproduction. The error analyses presented here portray a psychologically real representation of this model via performance generalizations. In turn, this model accounts for how the cognitive system executes heuristic operations across domains and levels in both a serial and parallel fashion, thus making it possible to explain clusters of multiple errors in the ASL-SRT task.

MATERIALS AND METHODS
THE AMERICAN SIGN LANGUAGE SENTENCE REPRODUCTION TASK (ASL-SRT)

The ASL-SRT was developed for sign language by adapting the approach used in the spoken-language Test of Adolescent Language 3 (TOAL3), Speaking/Grammar subtest (Hammill

et al., 1994). Like the TOAL3, this test presents sentences in gradually increasing complexity and asks the subject to repeat the sentence exactly. The 20^2 test items are graduated in difficulty, increasing in length of sentence, complexity of morphology, and number of propositions; **Table 1** lists word span, syntactic complexity, and content for each item. The first 10 test items are single clause sentences with a variety of argument-predicate relations, as shown in the top half of **Table 1**. In contrast, Items 11–20 contain multiple clauses with various types of relations among constituents.

The test is administered on a laptop computer. Subjects view a video of a woman who serves as both an instructor and a model producing the set of practice and test sentence items. She instructs subjects to copy the model's exact signing, stressing the need for verbatim response. This instruction is followed by three practice sentences with subjects responding. In the review of the practice items, the instructor compares two versions of the signs YESTERDAY and DARK for which she used one version in the practice session and presents the common alternate form, showing movement and handshape variants and instructing subjects to copy the exact parameters used by the signing model for each sentence. The test session follows and is self-paced without a time limit for response: subjects view each sentence only once, but they then have unlimited time to make their response. Thus, subjects may self-correct or repeat a response before moving on the next sentence by pressing a key. On average, it takes a subject 10 minutes to complete the test.

The responses were video-recorded and the rating took place later. In the case of repeated responses, raters were instructed to use the last response for rating purposes. In the absence of any response before moving on the next sentence, raters were instructed to mark the sentence item as a failure. On average, a complete rating of a subject's 20-response set takes 20 minutes.

RATER TRAINING

One compelling reason for pursuing this method for measuring ASL proficiency is that the test is easy to administer and can be scored with robust inter- and intra-rater reliability by native signers, even those without a linguistics background, following minimal training. Both rater training and scoring takes place with raters blind to the hearing status of the subjects. The training materials consist of a DVD, which includes training and practice videos in ASL, and downloadable rating sheets, scoring symbol keys and guidelines for the scoring of test sentences. Raters complete blind practice with sample training subjects drawn from a wide range of signing fluency, from novice to highly fluent. This enables them to develop metalinguistic skills for assessing a range of performance levels and familiarizes them with acceptable and unacceptable variation in sentence reproduction. It is not clear

---

[2] In Hauser et al. (2008) we described the development of a 39-item version of the ASL-SRT as well as its initial administration to populations with varied ASL backgrounds and skills. Initial administration of the test established the validity of the reproduction task as a tool for measuring fluency in a sign language, showing that the test does indeed discriminate among signers of varying ages and fluency. The current discussion focuses on a 20-item subset of the original 39 sentences. This subset of sentences comprises the second refined version of the ASL-SRT.

Table 1 | Word span and syntactic complexity of ASL-SRT items, with sentence content and inflections.

Item	Word span	Syntactic complexity	Sentence content and inflections
1	5	Transitive predication	INDEX-first FINISH BUY OLD HOUSE
2	3	Adjectival predication	THAT-i TREE TALL
3	4	Transitive predication	INDEX-i FINISH FIND KEY
4	6	Adjectival predication	MY LAST VACATION SEVEN YEARS AGO
5	4	Adjectival predication	THAT MAN NICE SWEET
6	4	Transitive predication	INDEX-i NOT LIKE INDEX-j
7	4	Adjectival predication	SUNDAY NEWSPAPER TEND CL: thickness-on-surface
8	4	Adjectival predication	MY DAUGHTER SELF-i AGE-THREE
9	4	Intransitive action	MY DOG CONTINUE+rep BARK
10	4	Adjectival predication	WOMAN SELF-i COMPETENT MATH
11	7	Copular object NP, adjectival predication	WASHINGTON #DC HAVE MANY GOVERNMENT BUILDING, CL: huge-object-alternating-ijk
12	4	Adverbial predication, intransitive action	INDEX-first DRIVE FIVE-HOUR, ARRIVE WORN-OUT
13	7	Conditional clause with transitive predication, consequence clause with adverbial predication	IF INDEX-i NOT BELIEVE INDEX-self, THAT FINE
14	4	Conjunction of intransitive action and locative predication	MOTORCYCLE CL: vehicle-slide-off-ground, HIT TREE
15	6	Locative predication, Transitive predication, Locative predication	WOMAN RIDE-horse HORSE, SEE-i FENCE, CL:jump-over-fence-i
16	6	Locative predication, intransitive action	THREE-OF-US GO-i-rep GRANDMOTHER HOUSE, HELP CLEAN-UP-arc-i
17	6	Locomotion, Locative predication, POV predication	INDEX-first LIKE GO BIKE PATH CL: trees-go-by
18	7	Transitive predication, Object complement, adjectival predication	#DAVID GO WATCH-i MAN LECTURE, CL: in-back-of-audience FULL
19	9	Transitive predication, transitive predication	SCIENCE TEACHER DISTRIBUTE TEST, INDEX-arc STUDENT HAVE-TO NAME+rep-on-list STAR
20	7	Locative predication, transitive predication	ONE LITTLE GIRL GO OUT, FLOWER CL: pick-up/ put-in-basket+rep-arc

whether non-native signers can be trained to achieve high accuracy in rating; we have focused on using native signers for this role, since it may be difficult for non-native signers to notice some of the errors that they might well make themselves.

This rater-training protocol provides an introduction to the overall accurate reproduction aim of the ASL-SRT and to the two aspects of rating: scoring each item as correctly or incorrectly reproduced, and noting the type of error in the case of incorrect reproduction at various levels. Raters also are introduced to the internal structure of the test, with sentences arranged in increasing levels of length and complexity. The rater-training tutorial proceeds through the following process: raters first build skills and familiarity with judging reproduction accuracy with the basic test sentences 1–10. They then proceed on to accuracy and error type notation for the more complex sentences 11–15 and 16–20, following a mid-point review of their skills with additional instruction. In general, the rater training takes about 3 days to complete.

SUBJECT POOL

For the analysis described here, subjects were recruited from deaf college programs and from summer camp programs for deaf children and hearing children with deaf parents. The subject pool was comprised of signers from three groups: Native Deaf adult signers (DDA), ages 15–30 ($N = 25$); native Deaf young signers (DDY) ages 10–14 ($N = 25$); and native Hearing adult signers (HDA), often known as Children of Deaf Adults, or CODAs, ages 15–30 ($N = 25$). The protocol was approved by the Institutional Review Boards of the University of Rochester and the Rochester Institute of Technology, and all subjects gave informed consent.

RESULTS

Seventy-five participants took the ASL-SRT and five trained raters rated the participants' sign reproductions independently. The inter-rater reliability was high and correlation coefficients ranged from 0.86 to 0.92.

For each participant, ASL-SRT performance was indexed by two different measures. First, reproduction was scored as correct or incorrect based on an all-or-none scheme whereby any error in reproduction would lead to a zero score. Second, more detailed analyses were carried out classifying errors by type and recording the frequency of each type of error within and between participant groups.

OVERALL RESPONSE ACCURACY ANALYSIS

Figure 1 shows the number of subjects in the 75-subject pool who accurately reproduced each of the test's 20 sentences. The slope

indicates an overall increasing difficulty in sentence reproduction, reflecting the increasing complexity of ASL grammatical structure from sentence 1 to 20.

The overall trend for each group for performance across the 20 sentences is shown in **Figure 2**. Grouping the subjects by similar home backgrounds but differing age and hearing status can ultimately help us to tease out which experiential factors may be responsible for the various fluency levels shown by the subjects.

An ANOVA was conducted with Group (DDA, DDY, HDA) as the between group factor and number of correct sentences reproduced as the dependent variable. A significant group difference was found, $F_{(2, 72)} = 16.001$, $p < 0.001$; partial eta squared = 0.308. *Post-hoc* analyses revealed no significant difference between the two deaf groups, DDA ($M = 14.7$; $SD = 2.8$) and DDY ($M = 13.7$; $SD = 3.2$). However, young and adult deaf signers were able to reproduce more ASL-SRT sentences than HDA ($M = 9.4$; $SD = 4.3$).

ERROR TYPE ANALYSES

The remaining analyses in this paper are based on the error ratings of the first author of this article who served as a rater for this 75-subject pool and the rating trainer. The analysis of sentence reproduction failures by the 75 subjects provides useful information on the trends for particular error types along the 20-sentence range of incremental complexity. For each incorrect reproduction response, the first three errors identified in the sequence of signs for each sentence were collected for quantitative analyses here (See **Table 2**).

Analysis of the first 3 errors as a methodological protocol captured the vast majority of errors and was well within and often beyond the average number of errors that signers made in a single sentence. At any given word location within a sentence, multiple errors were noted as well. The complete range of error types is explained and exemplified in **Table 3**.

The reproduction error types listed in **Table 3** do not include linguistic deviations reflecting factors such as dialect, age-related experiential differences, and permissible phonological variations in ASL. Rater tolerance to variation was resolved with tutorials where raters were exposed to 10 models including a mixture of novice, semi-fluent, and fluent signers. Rather, error types were incorrect reproductions and not merely pronunciation or accent differences.

In lexical and morpho-phonemic substitutions and morphological merging errors, subjects substituted a different sign than the one in the stimulus model in a given sentence location. Such data allow us to flesh out and further subdivide the notion of "semantic error" as described in Mayberry and Fischer (1989). In lexical and syntactic errors, we see a further distinction between errors preserving semantic content at the lexical level (synonyms) vs. errors preserving semantic content through grammatical alternations, affecting the morpho-syntactic structure of the entire sentence. In cases of multiple alternations or commissions in a given sentence location, each deviation is counted as a separate error.

In the next section we turn to the main factor affecting reproduction success and error type: the relative fluency of the signer. To begin, **Figure 3** shows the number of occurrences of six separate error types in the pooled 75-subject response data for 20 sentences. Each of the first 3 errors was included in the count,

FIGURE 1 | Histogram of participants with correct reproduction for each of the 20 sentences in the ASL-SRT task (N = 75 subjects).

FIGURE 2 | Number of participants per group (maximum N = 25) with correct sentence reproduction as a function of sentence complexity ordered from easiest (sentence 1) to hardest (sentence 20).

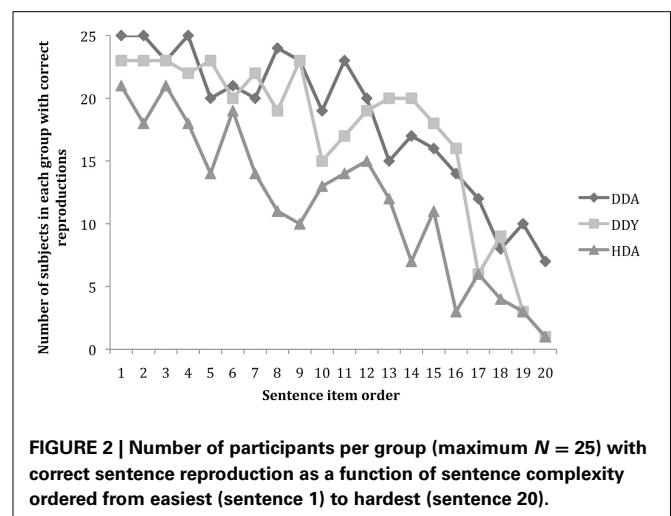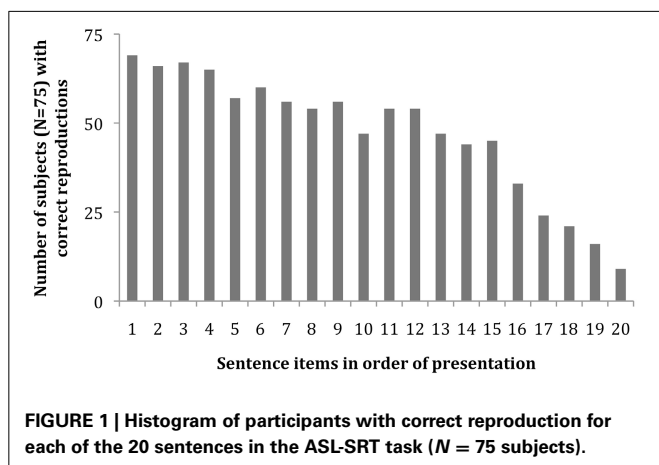Table 2 | Distribution of errors across 1500 responses.

Subject group	# test items	Successful reproduction	Failure	With 1 error	With 2 errors	With 3 errors or more
DDA (*n* = 25)	500	361	139	77	34	28
DDY (*n* = 25)	500	348	152	74	49	29
HDA (*n* = 25)	500	247	253	121	80	52
N = 75	1500	956	544	272	163	109

Table 3 | Classification of reproduction errors.

Error type and description	Example
OMISSION AS ONE TYPE OF ERROR	
Target sign is omitted	Stimulus: MOTORCYLE CL: vehicle-spin HIT TREE Response: MOTORCYCLE HIT TREE
MORPHOLOGICAL TYPE OF ERROR	
Bound inflectional morphology is replaced, resulting in simplified sign (morphological omission)	Stimulus: MY DOG CONTINUE+ + + BARK Response: MY DOG CONTINUE BARK (no reduplication for continuous aspect)
Re-interpretation of classifier structure such that response has similar form but different meaning	Stimulus: INDEX-first LIKE GO BIKE PATH CL: trees-go-by Response: INDEX-first LIKE GO BIKE PATH CL: wind-in-air
Merge two signs into one form	Stimulus: MOTORCYLE CL: vehicle-spin HIT TREE Response: MOTORCYCLE CL: vehicle-spin CL: vehicle-hit-tree
LEXICAL TYPE OF ERROR	
The target sign is replaced by a different lexical form (lexical substitution)	Stimulus: MOTORCYCLE CL: vehicle-spin HIT TREE Response: MAN CL: vehicle-spin HIT TREE
Sign not present in the stimulus item is added (lexical commission)	Stimulus: MOTORCYCLE CL: vehicle-spin HIT TREE Response: MOTORCYCLE CL: vehicle-spin HIT TREE CL: person-fall-from-vehicle
PHONOLOGICAL TYPE OF ERROR	
Response sign is misarticulated in form, thus recognized as different from the target sign (misarticulation)	Stimulus: MOTORCYCLE SPIN(base hand palm-down) HIT TREE Subject Error: MOTORCYCLE SPIN(base hand palm-up) HIT TREE
Sign is replaced with one that was morphologically/lexically related (morpho-phonemic substitution)	Stimulus: MY DOG CONTINUE+ + + BARK Response: MY DOG CONTINUE+ + + BITE
SYNTACTIC TYPE OF ERROR	
Sequence of signs is reordered (word displacement)	Stimulus: INDEX-first LIKE GO BIKE PATH CL: trees-go-by Response: INDEX-first LIKE GO CL: trees-go-by BIKE PATH
A sign is repeated at a different location in the sentence	Stimulus: INDEX-first LIKE GO BIKE PATH CL: trees-go-by Response: INDEX-first LIKE GO BIKE PATH CL: trees-go-by BIKE
OTHER RESPONSE ERROR TYPES	
Rough approximation of form and movement of target sign (visuo-motoric mimicry)	Overall response has a meaningless non-sign approximating the phonology of the target sign at a particular sentence location
Miscomprehension: Response indicates that subject does not understand concepts in stimulus	Stimulus: MOTORCYCLE SPIN HIT TREE Response: MOTORCYCLE RIDE SEE TREE

FIGURE 3 | Incidence and relative proportions of error types across 75 subjects.
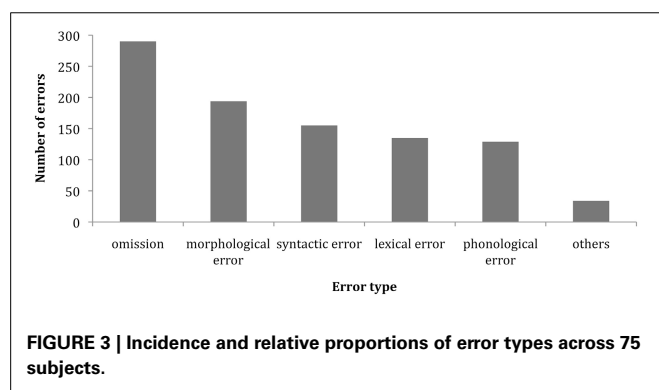
including duplicate types of errors within a particular subject response.

Word omission is the most frequently occurring type of error, with a higher incidence than the remaining error types: morphological, syntactic, lexical, and phonological. There is no statistically significant difference among the latter four error types in overall frequency of occurrence.

Error types as a function of hearing status and age

Our next step is to examine error patterns in the reproduction responses of signers as a function of their hearing status and age. Errors in signed responses are categorized by type and the proportion of each error type across the three groups in the subject pool is tallied. **Figure 4** shows the striking distinction in error type distributions across the three groups. Between the HDA group and the other two groups, there is a contrast in the most prevalent types of error, in that these hearing signers make more morphological, lexical and phonological errors than the two groups of deaf signers, whereas omissions and syntactic errors are comparable across all three groups. This seems to indicate distinctive differences in the respective groups' strategies for performing the task.

A multiple analysis of variance (MANOVA) was conducted to determine if there were differences in the frequency of error types (Omit, Morphology, Lexical, Phonological, Syntax) between three groups of native ASL participants (Deaf adults, Deaf youth, Hearing adults). Significant group differences were found for Morphological errors [$F_{(2, 72)} = 28.11$, $p < 0.001$, eta squared $= 0.44$], Lexical errors [$F_{(2, 72)} = 12.70$, $p < 0.001$, Partial eta

FIGURE 4 | Incidence and relative proportion of error types by hearing status and age.

FIGURE 5 | Incidence and proportion of error type by fluency level.
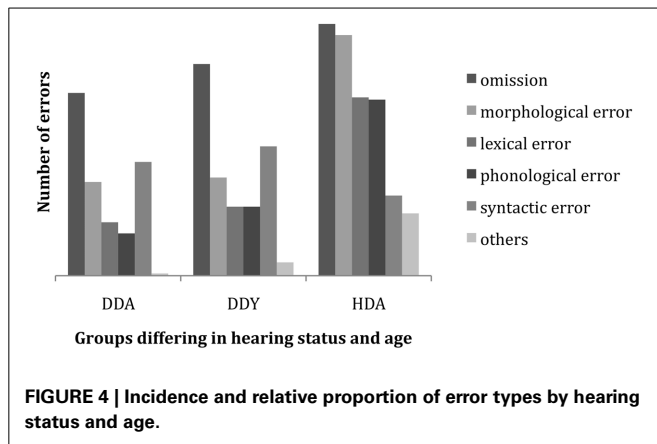
squared = 0.26], and Phonological errors [$F_{(2, 72)} = 11.11$, $p < 0.001$, Partial eta squared = 0.24]. There were no significant group differences for Omission or Syntactic errors. *Post-hoc* analyses with Bonferroni corrections revealed that Deaf adults and youth made fewer Morphological, Lexical and Phonological errors than Hearing adults. All were at $p < 0.001$ with the exception that the level of significance between Deaf youth and Hearing adults for Phonological errors was at $p < 0.01$. Deaf adults and youth had the same pattern of occurrence of errors across error types.

However, this is not the whole story. To understand the role and nature of cognitive mechanisms across levels of fluency, we re-grouped the subjects into high, moderate and low fluency groups. Then we investigated the strategies for the task within and across these groups as revealed by an in-depth error analysis. The results of this analysis are set out in the sections below.

Error types as a function of fluency

The 75 subjects were ranked based on their ASL-SRT performance as judged by the accuracy scores, and subjects with the top 25 highest correct reproductions were grouped as High (10 DDAs, 13 DDYs, 2 HDAs), the middle 25 as Moderate (11 DDAs, 6 DDYs, 8 HDAs), and the bottom 25 as Low (4 DDAs, 6 DDYs, 15 HDAs). The purpose of this re-grouping was to examine and describe the type of errors made by individuals of different levels of fluency across hearing status and age. In the figure below, we compare and contrast the error patterns of subjects who performed in the High (20–15 correct reproductions), Moderate (15–12 correct reproductions), or Low (12–2 correct reproductions) range.

Figure 5 below shows how the relative proportion of error types interacts with subject fluency, producing differing proportions of error types for signers at the high, moderate and low levels of ASL fluency.

A multiple analysis of variance (MANOVA) was conducted to determine if there were differences in the error types (Omission, Morphological, Lexical, Phonological, Syntactic) as a function of sign fluency (High, Moderate, Low). Significant fluency differences were found for all error types: Omission errors, $F_{(2, 72)} = 25.72$, $p < 0.001$, Partial eta squared = 0.42; Morphological errors, $F_{(2, 72)} = 19.79$, $p < 0.001$, eta squared = 0.36; Syntactic errors, $F_{(2, 72)} = 3.23$, $p < 0.01$, eta squared = 0.08; Lexical
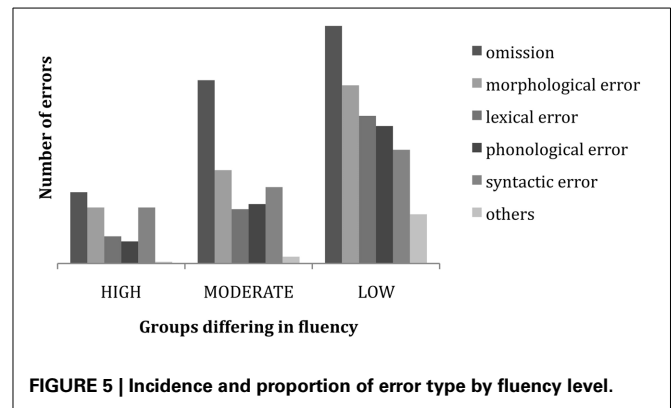
errors, $F_{(2, 72)} = 20.83$, $p < 0.001$, Partial eta squared = 0.26; and Phonological errors, $F_{(2, 72)} = 14.83$, $p < 0.001$, Partial eta squared = 0.30.

Post-hoc analyses with Bonferroni corrections revealed that the Low fluency group made more Omission errors ($p < 0.01$) and Syntactic errors ($p < 0.05$) than the High fluency group. In addition, the Low fluency group made significantly more Phonological errors than both the Moderate ($p < 0.01$) and the High ($p < 0.001$) fluency groups. The Low and Moderate fluency groups made significantly more Morphological and Lexical errors (both $p < 0.001$) than the High fluency group but did not differ from each other.

DIFFERENTIATING THE ERROR PATTERNS OF FLUENCY GROUPS

Our qualitative investigation of errors reveals that the overall structure of reproductions by fluent native signers under extreme task demand, while not perfectly correct, is nonetheless consistently well-formed, regardless of whether the signer is deaf or hearing. In contrast, among less fluent signers there is an increase in ungrammatical responses, with some target words omitted or replaced with unintelligible forms. This trend increases as the length and morpho-syntactic complexity of the sentences increase.

Even the first half of the ASL-SRT, made up of 10 sentences averaging four words long, led to reproduction failures among semi-fluent subjects. Although there was no increased length in words across these sentences, respondents experienced increasing difficulty due to the increase in morphological complexity. This is due to the fact that there are two types of stimulus items in the first half of the ASL-SRT. While all consist of a single clause, the first five test sentences contain only bare, uninflected lexemes, while the second five sentences have inflectional morphemes affixed to some of the lexical items. Some are aspectual inflections; others are nominal class markers (classifier morphemes). This difference produces an increase in the potential for bottleneck conditions through even the first half of the ASL-SRT.

The determination of what causes the complexity of reproduction in the second half of the ASL-SRT test is less straightforward. In the construction of the test, the items were ordered empirically, based on how accurately subjects responded. For these most complex 10 sentences, verbatim serial memory obviously becomes more challenging as the sentence becomes longer. While these

items do involve multiple clauses, it is not clear what precise features of structural complexity contribute to the psycholinguistic complexity of the reproductions.

With this perspective on such cognitive bottlenecks, we would expect to see a boost in performance accuracy if the subject had an opportunity for rehearsal using a variety of heuristics to reproduce the individual word components or the overall sign sequence of a test sentence. There is, however, no opportunity for rehearsal in the ASL-SRT. The subject only has one chance to see the test sentence and must work through potential bottlenecks in processing to reproduce even very complex, lengthy sentences with only the resources on hand.

In a sense, our data support a straightforward insight: we would not expect semi-fluent signers to be able to use non-linear scaffolding in their working memory in the way that fluent signers can. Whatever kind of working memory semi-fluent subjects have will determine the quality of their performance on the ASL-SRT task. Within this group, potential misarticulations might be replaced with unintelligible forms whenever subjects are overwhelmed by the task. In this context, lexical misarticulation by fluent signers will often include replacement of particular features. In contrast, we would expect an unintelligible form by a non-signer to be articulated with no constraints on the linear segmentation or inflectional prosody. In the case of a relative lack of knowledge of sign language phonology and morphology, visuo-motoric imagery may be a useful *ad-hoc* solution when attempting to process and imitate a string of linguistic word forms presented in sign language.

In this sense, the ASL-SRT task is significantly different from list recall. In a list recall task, subjects' reproductions are limited by the number of words in the list and by classic phenomena such as primacy and recency. As in spoken languages, once an individual word has a particular function in relation to other words in the sentential sequence, this affects how this word is encoded and then reproduced. In short, then, a mechanism other than serial word-list memory is required to explain the error patterns which we find across our subjects.

Here, we may draw on recent research on sentence reproduction processing. A distinction in error types revealing automatic vs. effortful processing has been modeled by Ronnberg et al. (2008) as implicit vs. explicit processing. This psycholinguistic model of online processing and remembering highlights the efficiency of implicit processing. In related literature, Potter and Lombardi (1990) have suggested a model of verbatim recall by native, fluent language users which relies on conceptual storage and regeneration of language structure in memory. In their model, verbatim recall of sentences relies on recent lexical activation.

RESULTS OF ERROR ANALYSES

Fluent Deaf and Hearing signers obtained higher reproduction accuracy scores and made different sorts of errors than weak signers. The reproductions of highly-fluent subjects among the DDA, DDY and HDA groups often differed from the stimulus in lexical ways, while retaining and faithfully repeating underlying aspectual and other sub-lexical morphology. In contrast, weaker signers committed a greater number of morphological errors and

omissions. Another distinctive profile among weak signers is that the number of well-formed words they produce (no matter if correct or not) remains constant throughout the test segments (which increase in complexity), revealing a specific cognitive limitation where grammatical knowledge is lacking. Also, any adult or child signer could potentially exploit visual and motoric imitation to overcome lexical and morphological limits, and this can lead to a superficial illusion of comprehensible signing. However, while we see such misarticulations among native signers across the range of fluency, these vary in their degree of grammaticality, with some showing a high formal resemblance to the target and others clearly non-signs. In this sense, the data reveal that signers' error types group according to individuals' relative level of competence and knowledge of ASL grammar. Thus, we would claim that the structure of the subjects' responses projects the affordance and constraints of deep linguistic representations to differing degrees, and subjects resort to alternate processing strategies in the absence of such knowledge or under conditions of high task demand. We will lay out observational generalizations highlighting these points and then provide interpretation of the data in support of the theoretical models cited above.

GENERALIZATION 1: TENDENCY TOWARD SIMPLIFICATION FOR PARTICULAR WORD CLASSES

The ASL-SRT task requires the subject to reproduce peripheral details along with main propositions. Occasionally a subject will eliminate peripheral details, especially determiners and qualifiers, when reproducing a target sentence. For example, the DET class in the subject NP position throughout the test is a construction prone to error. This is consistent both within and across subjects.

For certain word types, there are also constraints on the types of replacement errors produced, which tend to stay within the target class. For example, replacements for determiners stay within the class of determiners. The test items involve two different types of DET, but common replacements for both are the more generic INDEX or the target DET item without its spatial agreement inflection (indicated by 'i'). Item #2 contains THATi (THAT with a spatial agreement inflection); items #8 and #10 contain SELF+locus-i (self with a spatial agreement inflection). In all of these cases, the target item is replaced by a less-marked DET of the same kind. It is rare that a more highly-marked or inflected DET form replaces a less-marked or uninflected form.

It appears that the surrounding context of a particular word can trigger a process that results in omitting or replacing a word or particle morpheme. This happens more often in the second half of the ASL-SRT (Items #11–20), where there are more signs to be reproduced and greater linguistic complexity of the sentences. As we will see in other examples, it appears that the sentence structure and task create a possible chain of errors in which the subject mistakenly encodes (or fails to encode) a definite or specified NP construction, thus taking a wrong turn in the processing of the sentence. Incomplete interpretation of the sentence can create such missteps in processing regarding a particular noun argument or its relation to other noun arguments, and the interdependence of grammatical operations can then lead to additional errors in the sentence.

GENERALIZATION 2: INTERDEPENDENCE OF MORPHO-SYNTACTIC ERRORS

Other reproductions of items containing DETs show that the position of the DET/specifier may shift, the DET may be omitted, or the DET may be copied to the beginning or end of the determiner phrase or of the entire clause. This can be seen in the reproductions of Sentence #20, where the determiner ONE appears beside the adjective LITTLE and the noun GIRL: ONE LITTLE GIRL. Two subject responses are: Target: ONE LITTLE GIRL vs. Response: GIRL LITTLE ONE or GIRL LITTLE. In the first response, the word order deviation can be viewed as a pragmatic variant, since bracketing of a phrase by a repeated determiner is a common ASL device for focus or emphasis; and prenominal adjectives are more frequently displaced after the noun rather than to any other position in the sentence. Alternatively, perhaps the subject initially omitted DET and ADJ by mistake and then filled in the omitted material afterwards. But in either case, the displacement is constrained, with the DET omitted or displaced to a position after the clause.

Omission of DET occurs most often among the subjects we tested and thus appears to be a common response to serial memory limitations during the reproduction task. In contrast, omission or misplacement of the head noun GIRL is rare, presumably due to its syntactic salience and to the fact that the adjacent words ONE and LITTLE depend on its appearance. Overall there is a hierarchical relationship among these three words, with their role in the phrase determining the likelihood of their appearance and position in responses. These data support a constraint-based theory of reproduction performance. Other classes of words (modals, qualifiers or quantifiers) follow a similar pattern.

GENERALIZATION 3: PROCESSING CHOKEPOINTS

In our analysis of sentence responses, we also identified specific intra-sentential locations where errors were likely to occur across all groups. We call these locations chokepoints: sentence locations where processing bottlenecks occur, as indicated by a high frequency of reproduction errors at that point in the sentence. However, the type and extent of errors in and beyond this point in the sentence were likely to be quite varied. The type of error resulting from a particular chokepoint depends on two factors:

(1) the general fluency of the signer, and (2) lexico-morpho-syntactic complexity of a particular word in a sentence. The latter factor can induce a series of bottlenecks for a particular sentence item. Beyond this slot in the sentence, additional error types and number tend to cluster for signers, suggesting a non-linear hierarchy of grammatical domains constraining reproduction in these challenging conditions. The effects on a particular word can come from its visual, semantic or syntactic resemblances with different words in the lexicon or from its long-distance grammatical relations with other words in the sentence.

These chokepoints are not limited to a single grammatical domain. Earlier we illustrated the errors occurring within the grammatical domain of Determiner, for sentence items #2, #8, #10, and #20. In contrast, the errors for Sentence Item #4 as shown in **Table 4** occur at several chokepoints where time and number signs occur.

The error distribution in **Table 4** shows an overall pattern like that for the adjective LAST as one primary chokepoint in the sentence. This slot in the sentence shows a greater number of errors occurring in comparison to the other items in the sentence. In terms of the subjects' thinking, the second word LAST may be confused with AGO; or some subjects may have only the sign AGO for expressing the meaning LAST. Such errors validate the concept of a sentence framework in which sentence slots that occur in a common linguistic frame may allow a swapping of words with a similar function or "spread" of the same word to a slot with a similar function. In these cases, we see the phenomenon detailed in Potter and Lombardi (1990), where a given conceptual content triggers a particular syntactic frame to be regenerated without verbatim recall.

Another chokepoint in this sentence is the numeral sign SEVEN. Subjects may misperceive how many fingers are extended, often replacing 7 with the numeral 3. Nevertheless, the replacement is still a numeral, a fact which demonstrates the constraints that their grammatical knowledge places upon their errors. Also in connection with this slot in the sentence, some subjects overgeneralize an ASL rule which allows for incorporating a numeral handshape into the following sign AGO, generally up to the number 5. Spreading the handshape for the numeral 7 throughout the following sign, YEARS-AGO, violates

Table 4 | Frequency and type of error made for sentence item 4.

Rank	Group	MY	LAST	VACATION	SEVEN	YEAR	AGO
ONE ERROR							
19	DDY				Insert AGO (syntactic)		
48	DDY				Replace w/3 (phonological)		
65	HDA		Replace w/AGO (lexical)				
71	HDA		Replace w/AGO (lexical)				
73	HDA				Merge 7-YEAR (morphological)		
TWO ERRORS							
38	HDA		Replace w/AGO (lexical)	Replace w/CLASS (lexical)			
68	DDY		Replace w/AGO (lexical)		Replace w/3 (phonological)		
THREE ERRORS							
55	HDA	Omit	Replace w/AGO (lexical)	Omit			
75	HDA		Omit	Omit			Omit

this combinatorial constraint, yet at the same time reveals greater knowledge of ASL structure.

In response to this sentence and in contrast to the types of errors above, some subjects produced a series of unintelligible forms. In some cases they placed their hand on their shoulder, thus indicating they noticed the articulation of the AGO segment. However, their choice of handshape was wrong, resulting in a non-sign. One young native signer misarticulated this sign by placing his hand on the opposite shoulder.

GENERALIZATION 4: CO-DEPENDENCE AMONG GRAMMATICAL OPERATIONS REVEALED IN A RANGE OF SURFACE OUTCOMES

We now turn to examples of error patterns illustrating the interdependence of multiple processing strategies in the various grammatical domains of ASL. The sample analyses of responses here reveal the interdependent relation between misarticulation, omission and displacement made by subjects around chokepoints within individual sentence items. In ASL, assimilation across morphemic segments can result in a complex non-concatenative form. However, the fluent signer may cognitively parse these as separate morphemes during encoding of the stimulus sentence. This grammatical knowledge may then result in specific sorts of errors. Moreover, the resulting omission of a word may impact the well-formedness of the overall sentence response. However, the option for omitting a word is constrained by the grammar.

As described earlier, our analyses reveal a correlation between the grammaticality of the overall response and the fluency of the subject. The more accurate the subject was in performing the whole task, the more likely an omission is to be triggered by a grammatical operation generating an acceptable alternate sentence form. Sentence Item 6 is an example with a morphosyntactic condition, which may trigger such top-down processing errors. In the target sentence, the negator NOT is separate from the verb LIKE. Subjects often merge NOT LIKE, preferring the alternate ASL contracted form. Some even produce double negation, in which they produce both the sequential NOT LIKE and the contracted form in the same sentence, a violation of the rule of negation in ASL.

One explanation of such errors is the independence of the negator in the modal domain vs. in its bound form in the verb-internal domain. A subject can fail to coordinate the two domains when functioning under bottleneck-inducing conditions. Furthermore the negative contracted verb may be encoded as a single lexeme and thus reproduced independently of other linguistic forms. This can lead to the redundant outcome.

At other times, the range of morphosyntactic commission choices mirrors the range of possible word replacements although certain replacements may be triggered by semantic and syntactic motivating factors, either within the local phrase structure or across multiple phrases. In contrast to the chokepoints in Sentence #4, which involve the grammatical domains of time and number, the errors in Sentence #7 as shown in **Table 5** involve chokepoints which relate to linking/copular verbs and size-and-shape specifying classifier predicates.

Such output may still show effects seen in STM, such as primacy or recency effects. Also among semi-fluent signers, we have found a larger proportion of visuo-motorically driven formation

along with some linguistic fragments from ASL phonology. This shows that even a weak exposure and fluency level in ASL results in performance constraints of a grammatical nature rather than pure visual perception or imagery.

GENERALIZATION 5: REVEALING RELATIVE GRAMMATICAL PROFICIENCY VIA A LEFT-TO-RIGHT CHAIN OF ERRORS

The rigorously-controlled ASL-SRT protocol helps to reveal behavioral patterns in the manual-visual modality among those who have had minimal opportunity for learning or experience to acquire genuine and complete linguistic encoding. Error trends among semi-fluent subjects suggest a kind of scaffolding mechanism that relies more on episodic memory, a type of memory that encodes experiences that are rich with temporal, visuo-spatial, and emotional information. As a result, when they respond to complex test items in the second half of the ASL-SRT, they often reproduce only 3–4 actual words and resort to unintelligible formation or omission for the rest of the items in the target sentence.

Even so, this behavior is still constrained by grammar. Some semi-fluent signers produced unintelligible attempts or omission of less familiar words throughout the sentence and reproduced familiar words accurately while maintaining the overall word order. Other subjects started recalling the sequence of familiar words in a row, as if they were maintaining the order of their appearance in the test stimulus. For the remainder of the sentence, they ended their response with unintelligible forms.

It is essential to note that in all but the least-fluent signers, alternative options for sentence reproduction are constrained by grammatical boundaries for binding linguistic elements. The distance for displacement of a given word in a sentence, for example, is the result of a series of serial and parallel processing decisions. Such a "chain-reaction" phenomenon for reproduction is constrained by clause-internal restructuring as well as by the extent of the bottleneck and the increasingly severe types of errors it induces, such as the descent from local misarticulation into phrasal unintelligibility. Such interfaces are more complex in the test items involving multiple clauses. The error examples in **Figures 6–8** below are extracted from responses which are typical across all but the least-fluent signing subjects. The errors can show up in a variety of sentence response contexts, from a single isolated error to a series of related errors triggered by choices in sentence recomposition. As an example, the errors among adjacent words in Sentence Item #15, pictured in **Figure 6**, reveal several kinds of cascading interactions among the multiple operations for constructing poly-componential predicates in ASL.

In this response error, the signer bracketed the noun FENCE with the verb JUMP, once in a plain form and once with a locative form of the verb. In this case, the subject was able to merge the nominal class marker into the last verb as well, thus apparently introducing a serial verb construction (Fischer and Janis, 1990; Supalla, 1990). In other examples, the last verb is missing.

When adjacent words are merged in prosodic assimilation, we might assume the non-linear coalescence will reduce cognitive load, thus helping with the on-line processing of the sentence (Liddell and Johnson, 1989; Brentari, 1998; Sandler and Lillo-Martin, 2006). We often see such natural spreading

Table 5 | Frequency and type of error made for sentence item 7.

Rank	Group	SUNDAY	NEWSPAPER	TEND	CL: thickness-on-surface
ONE ERROR					
42	DDA				Replace primary CL (morphological)
48	HDA				Replace second CL (morphological)
52	DDA			Replace w/SENSITIVE (lexical)	
56	DDY	Omit			
59	HDA				Misarticulate w/CHURCH (phonological)
61	DDY				Replace second CL (morphological)
62	DDA			Omit	
66	HDA			Replace w/SOMETHING (lexical)	
TWO ERRORS					
58	HDA		Replace w/PAPER (lexical)		Replace primary CL (morphological)
69	HDA			Omit	Omit
70	HDA				Replace w/CHURCH (lexical)
71	HDA			Omit	Omit second CL (morphological)
72	HDA			Omit	Omit
74	HDA			Replace w/SENSITIVE (lexical)	Replace w/CHURCH (lexical)
75	HDA			Replace w/LIKE (lexical)	Replace w/CHURCH (lexical)
57	HDA			Replace w/SENSITIVE (lexical)	Replace primary CL (morphological) omit second CL (morphological)
68	DDY	Replace w/SATURDAY (lexical)		Replace w/NOW (lexical)	Misarticulate w/CHURCH (phonological)
THREE ERRORS					
73	HDA			Insert INDEX (syntactic) replace w/PREFER (lexical)	Omit second CL (morphological)
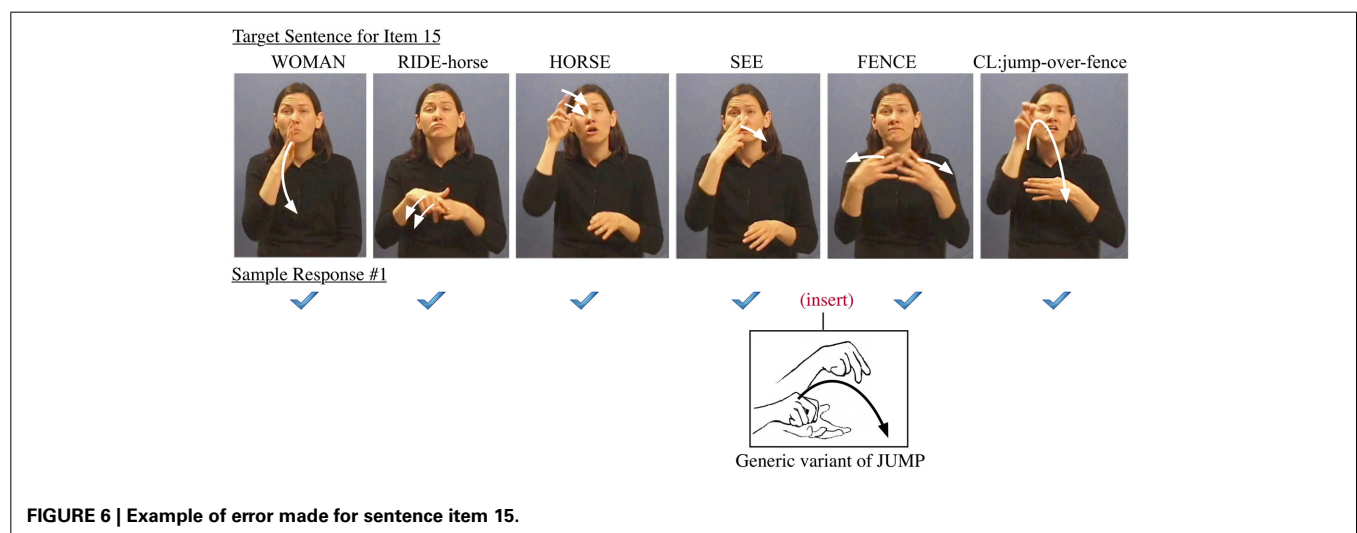
FIGURE 6 | Example of error made for sentence item 15.

of, for example, Weak Hand features to adjacent words. This is seen, however, only in specific morpho-syntactic contexts for fluent signers. For example, Sentence Item #15 has the prosodic scope of the spread Weak Hand feature extending from the verb RIDE to two subsequent words HORSE and SEE (see how

the weak hand is maintained in the second and third photo in **Figure 6**).

Fluent signers do not extend such prosodic assimilation across phrasal boundaries into the sequence FENCE JUMP which involves poly-morphemic classifier constructions. In contrast,

less-fluent signers do often assimilate in violation of the boundaries of words. This phonological assimilation can contribute to cascading errors, where less fluent signers may carry over the Weak Hand feature from RIDE to the last verb JUMP (see **Figure 7**). This correlates with their failure to merge the nominal class marker of FENCE into the last verb, since the Weak Hand is already occupied by the spreading Weak Hand feature from the earlier sign, as seen in the error below.

GENERALIZATION 6: REVEALING GRAMMATICALLY CONSTRAINED COMMISSIONS

If comprehensible articulation is achieved, there is still a gradient of accuracy in word reproduction. Each error is either phonologically or syntactically constrained by available options. In other words, the signer selects specific linguistic properties for matching the target form. The choice of alternate features is likely to be formally constrained by the phonology when the subject making this sort of error is a native signer with adequate fluency. If sufficiently varying features make it clear that a different lexical representation is involved, then the error is identified as a lexical commission (i.e., word replacement for RIDE-horse with a generic variant of RIDE). Furthermore, in accordance with the syntactic operation triggering this feature spreading, the woman must be considered as the subject of the verb JUMP (and hence as the subject of the generic RIDE variant). The agrammatical response ("jump out of conveyance and leap into the air") can be viewed as a phonological error, which is a consequence of the interpretation of the first verb in the target sentence.

From these sorts of errors, we see that success in reproducing Item #15 requires clausal scope for cognitive planning to preserve noun and verb relations. Moreover, Sentence #15 has an additional challenge, as the same hand configuration appears in several verbs throughout the sentence, with each use referring to a different noun argument. Such similarity in hand configuration can mislead some signers about noun relations. Evidence for this occurs in the errors of subjects who constructed responses in which the horse was the subject of the entire clause. Other subjects misunderstood the sentence in a different way, using the first subject WOMAN as the agent for jumping over the fence. Such syntactic errors are clearly grammatically constrained.

In the reproduction of sentence #17 on the ASL-SRT, there was a wide diversity in sequences of locomotion and path predicates. **Figure 8** illustrates three sample responses to Item 17. The first example has the pointer morpheme merged with the path morpheme displaced from a subsequent word, leading to a morphological commission error in the outcome.

Wherever the original meaning was maintained while the target form was replaced, a particular response could be treated as acceptable for ASL (though not correct in the context of the ASL-SRT), as in the second example. As with other constrained error examples, the deep structure is the same while the surface structure reflects a different output as a result of an alternate combination of multiple target morphemes. Here a separate nominal lexeme TREE was embedded into the complex predicate TREES-GO-BY, resulting in a double bracketing of the predicate.

The third response example illustrates how visual cues may affect the subject's encoding of complex sentence stimuli. The lexical replacement WIND, and the subsequent need to insert a different sign following it, is a "chain reaction" effect and an unacceptable error. If we compare the three error examples, the first and second are acceptable (though not accurate), but the third is different, since the subject apparently reconstructs a structure from a partial short-term memory of the original stimulus. Here a series of deviating nominal and verbal morphemes were put together, resulting in a meaningful and grammatical phrase. However, this outcome was not a rephrase of the target sentence. In establishing categories of "good" vs. "bad" errors, we suggest that, for each word in a sentence item, there was a range of possible grammatical and agrammatical deviations from accurate reproduction.

DISCUSSION, CONCLUSIONS, AND IMPLICATIONS

In our analyses thus far, we have discovered that overall accuracy on the ASL-SRT can be predicted by the hearing status and age of the signer. However, the best predictor of error types is the overall fluency of the signer. That is, fluent deaf and hearing

Target Sentence for Item 15
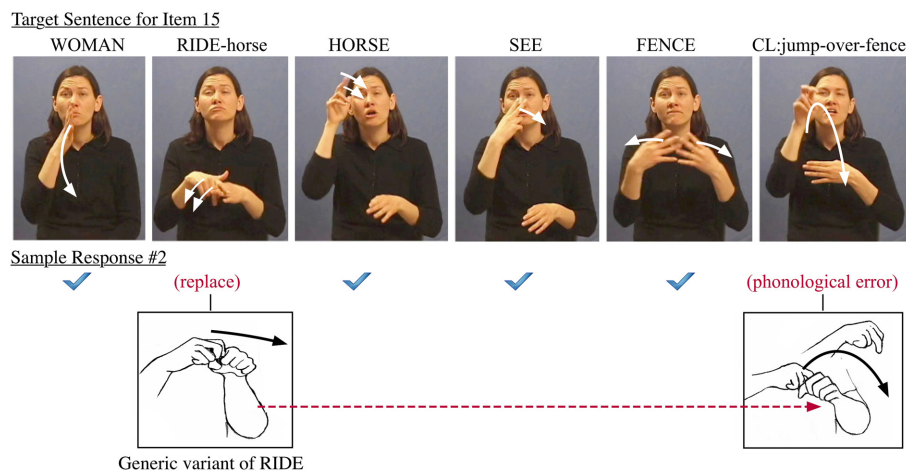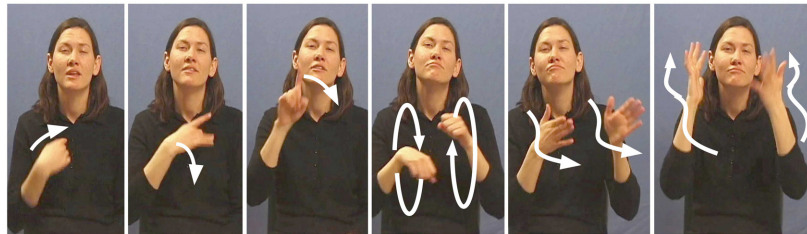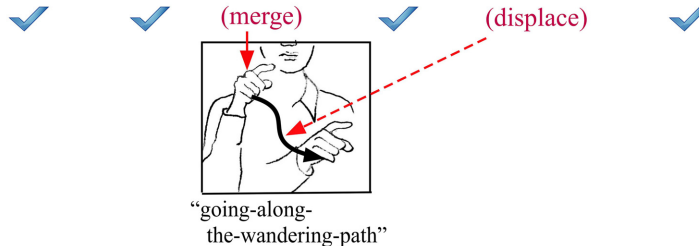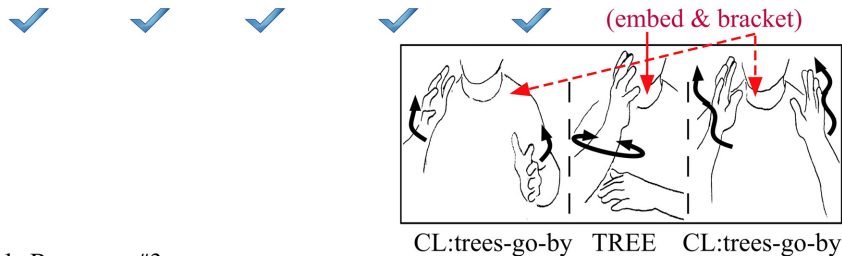WOMAN   RIDE-horse   HORSE   SEE   FENCE   CL:jump-over-fence

Sample Response #2
✓   (replace)   ✓   ✓   ✓   (phonological error)

Generic variant of RIDE

FIGURE 7 | Another example of error made for sentence item 15.
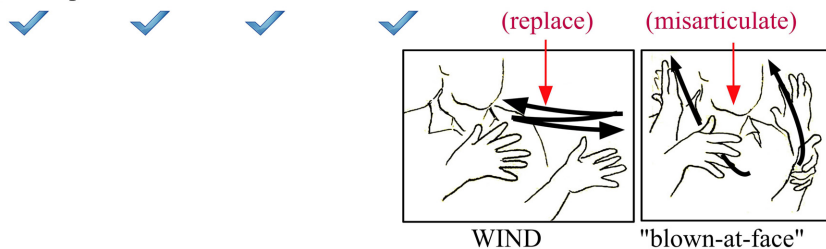
FIGURE 8 | Sample errors in responses to Item 17.

signers differed from less fluent signers in the proportion of differing error types. **Figure 5**, left, represents the pattern of error types and their frequencies made by signers achieving scores in the top third of all signers on successful reproductions. At times, these data show some likelihood of either omitting a word or producing a grammatical alternate form. But overall these subjects maintain a high level of accuracy in sentence reproduction. The main distinction between those representing the middle third (moderately-fluent signers) and the least-fluent third is the choice of strategy for processing and performing our increasingly complex reproduction task. While these two groups generate similar numbers of errors, the moderately-fluent signers seem simply to amplify the error trends of more fluent signers. That is, they are more likely to omit a word or create a grammatical error involving either a morphological or syntactic alternation.

This supports the Potter and Lombardi (1990) claim that verbatim recall is due to recently activated lexical items coalescing into a coherent and rich conceptual trace for the sentence. In weaker subjects, there may not be a mental representation of particular lexical items, preventing verbatim recall of the normal type. Less fluent signers are instead more likely to misarticulate or replace words. The basic profile of lexical errors in ASL-SRT performance across fluency levels and sentence complexity reflects both lexical error commission and unintelligible misarticulations, with the latter increasing as fluency declines. Among weak signers the distributional pattern of lexical omissions, commissions and displacements is least predictable. These subjects are more likely to produce unintelligible forms, as if they are attempting to match the target form through visual-motoric imitation with no idea of what the word means. The criterion we used to distinguish an

incorrect lexical item from an unintelligible form is the recognizability to the rater of the lexical root on which the misarticulation is applied. Careful investigation of the bar graphs in **Figures 4**, **5** reveals an increase in unintelligible forms (labeled as "other type of error") as fluency decreases and an increase in alternate sign forms (often categorized as morphological or syntactic error) as fluency increases among native signers.

In this analysis, for highly fluent signers there is no random noise in the data, but instead a strongly constrained performance. What this may indicate is that once a subject achieves complete fluency, rapid processing and deep-structure grammatical/semantic coding is available as a top-down scaffolding route for working memory. Such cognitive bootstrapping from deep structure processing serves them well in the end, producing their top-end-skewed performance curve for the 20 sentence items in the ASL-SRT (See **Figure 2**). It seems that younger signers may not have yet achieved the far end of this curve, as seen by their drop in performance for the last few (most complex) test items.

In the ASL-SRT task, we hypothesize that the working memory performance is based on content-addressable memory structures, and not on ordered phonological representations like those used in the recall of random lists (Potter, 1990; Potter and Lombardi, 1990). Thus, the type of order information necessary during sentence reproduction processing is considered to be different from the slow, temporal order processing that mediates list recall (McElree et al., 2003; Lewis et al., 2006). This process is akin to the implicit processing incorporated in the Ease of Language Understanding (ELU) model of working memory by Ronnberg et al. (2008) and the conceptual regeneration process outlined in Potter and Lombardi (1990). The importance of conceptual representations for STM of scenes and sentences has been established in these works and in Haarmann et al. (2003).

In this sense, highly fluent signers' inaccurate responses can be partially attributed to paraphrase guided by deep structure processing, leading them to produce a cascade of lexical and morphosyntactic changes when, for example, the choice of an equivalent lexical item leads to an additional difference in the order of signs. This structured type of grammatical variation requires an architecture for coordinating multiple layers of linguistic processing for sentence decoding and recomposition. This likely involves the interaction of clause, phrase, and word levels, with the integration of features from different tiers of information orchestrated by an overarching representation of the meaning and structure of the sentence. In other words, the conventional model of serial processing for non-sentence material is here replaced by a hierarchical model with parallel processing capabilities, a top-down scaffolding mechanism that assists sentence reproduction. For subjects at different levels of fluency, there appear to be some important psycholinguistic differences:

1. For subjects with a low level of fluency, the encoding bias is toward a visuo-spatial strategy in which the surface physiological features of the hand configuration, handshape and hand movement trajectory are copied from the target. This often results in an unintelligible response, with no recognizable signs or grammatical features.

At other times, the response may be more grammatical, but it will often involve multiple errors, each of a different type, independent of the others. Some signs may be omitted while others are misarticulated.

2. For moderately-fluent subjects, the bias is toward a linear strategy where individual signs and their syntagmatic positions are recognized and stored. The responses feature more accurate reproduction of the signs and their temporal sequencing in the sentence. Internal morphology is often deleted, and specific individual signs may be semantically replaced and/or phonologically misarticulated.

3. For highly-fluent subjects, the bias is toward a rapid processing of the semantic content of the sentence and a re-construction of the surface sentence composition guided by the deep structure grammar of ASL. Interestingly, this may put fluent native signers at a disadvantage in exact reproduction tasks. This often results in non-exact reproduction, as alternate grammatical forms for the stored meaning can be chosen and a shift in one grammatical element can cause cascading changes (scored as errors in an exact reproduction task) elsewhere in the sentence. This tendency increases with item difficulty. As a consequence, morpho-syntactic combinatory constraints often cause commission errors in the response. In this sense, while the test can function as a screening instrument for overall proficiency levels, its full value as a diagnostic instrument is realized with additional error analysis of responses.

In differentiating unintelligible articulation and constraint-based deviations at particular points of high task pressure, or "bottlenecks," it is likely that the signing of an unintelligible response reflects a certain limit in working memory capacity, where the misarticulated form corresponds to the collapse of linguistic encoding. In contrast, more fluent signers may simply display errors of lexical commission, reproducing alternate morpho-syntactic configurations because they have been able to process sentences more deeply and rephrase the words to sustain the sentence meaning. **Table 6** lays out our hypotheses about these on-line processing heuristics.

The performance generalizations articulated above portray a psychological representation of this model. The escalating demands of the reproduction task result in clusters of various

Table 6 | Modeling the correlation of error type to the layering of grammar.

Type of processing	Type of error	Domain of grammar
Top-down	Syntactic re-phrase	Deep-structure and semantics
	Morpho-syntactic alternation	Syntactic inflection
	Syntactic displacement or reversal	Word order
	Morphological omission or alternation	Lexical inflection
	Lexical omission or commission	Lexical formation
Bottom-up	Lexical misarticulation or unintelligible form	Sub-lexical encoding

types of errors, which are useful for teasing out processing at the interface between the layers of processing and specific grammatical domains. In turn, this model accounts for how the cognitive system executes heuristic operations across domains and levels in both a serial and parallel fashion, thus making it possible to explain clusters of multiple errors in the ASL-SRT task.

The generalizations outlined above are consistent with several current models of general language processing. First, we see clear evidence for the model put forth in Potter and Lombardi (1990) of regeneration of conceptual content in accordance with grammatical constraints and prompted by recent lexical activations for verbatim recall. In highly fluent signers, we see cascading interactions among the multiple operations for constructing poly-componential predicates in ASL. In contrast, semi-fluent signers exhibit isolated error patterns when multiple errors in a single item are seen, indicating a lack of adequate conceptual understanding to create grammatical regeneration. Such errors also provide support for the Ronnberg et al. (2008) model of effortful "explicit" processing of individual lexical items, without the time to build the entire sentence through this process. Second, the retention of morphological concepts across sentence items in fluent signers indicates sentence comprehension and the formation of a sentence composition plan for a response, with working memory making use of the grammatical architecture to link morphemic constituents.

The inclusion of subject groups who vary in fluency levels has added rich data to the testing of such models of language processing. The intuitive distinction between "good" and "bad" errors reflects a sense of different types of cognitive organization across fluency levels. The coordination of individual linguistic operations to accomplish the reproduction task suffers as fluency decreases and task difficulty increases. For this analysis, we have posited three processing strategies in use by signers: top-down linguistic analysis, linear processing at the individual sign level, and visual-motoric "copying" of the stimulus. Each of these strategies points to a particular interaction between signer fluency and cognitive skills in accomplishing the reproduction task. We can imagine a hypothetical efficiency trajectory for each scaffolding strategy in sentence reproduction throughout the ASL-SRT task. Each strategy will peak at a particular point in the increasing complexity of potential bottleneck-inducing stimuli. In order to achieve further proficiency, a signer would need to switch from "episodic" to "linear" and finally to a "non-linear" type of scaffolding. Each of the strategies outlined above fits well within the models mentioned. These three strategies are: first, a strategy of visuo-motor episodic mimicry among semi-fluent signers; second, an explicit lexico-syntactic processing strategy where serial order is maintained; and third, a faithful top-down re-generation of sentence composition.

In episodic mimicry, we see an attempt to process language without a foundation for either explicit processing or conceptual regeneration. In the lexico-syntactic processing heuristic we see access to recent lexical activation without full or timely conceptual processing skills. In the reproduction of fluent signers, we see a range of possible "chain of error" outcomes, which may deviate from the stimulus for complex sentences. This indicates the availability of linguistic scaffolding and

parsing options during cognitive and linguistic encoding and production.

CONCLUSIONS AND IMPLICATIONS

The ASL-SRT test paradigm, with its increasing complexity and bottleneck conditions inducing errors in reproduction, reveals distinctive cognitive strategies across signers varying in fluency while controlling for language background. The specific details of a signer's experience with ASL in the home can apparently create the conditions for a particular heuristic strategy to be employed as part of that individual's available scaffolding and approach in coping with a stimulus item. This points to a range of cognitive strategies in working memory for the visual-gestural mode, which then interact with formal constraints of grammar to support the top-down processing capacity that fluent native signers possess.

While the data in the present study were all collected from native signers, there are similarities between what we have found as error types in our less fluent signers and error types that were found by Mayberry and Fischer (1989) in their study of sentence shadowing by native and late learners of ASL. A number of investigators have shown that late learners of ASL typically achieve lower levels of ASL fluency, even after full immersion and many years of language use (Mayberry and Fischer, 1989; Newport, 1990; Mayberry, 2010). As discussed earlier in this paper, Mayberry and Fischer's (1989) shadowing results showed that native signers' errors were predominantly semantic: they correctly represented the meaning of the target sentences, though sometimes changing the structure as they shadowed. In contrast, late learners' errors were predominantly phonological. This pattern is strikingly similar to the tendency of highly fluent signers in the present study to retain the deep structure of target sentences, whereas less fluent signers made a variety of more superficial errors and changes. Unfortunately we cannot discern without further analysis whether the representational and processing strategies of late learners are precisely the same as those of less fluent native signers, but these similarities in error types suggest that this may be the case.

The fact that this cognitive approach encompasses both the spoken and signed processing of language is noteworthy. Errors in the sentence reproduction task follow similar constraints as errors in natural language production. For example, lexical commissions usually respect word category, and misarticulations are constrained to possible word formation. Furthermore, this kind of data analysis has proven essential in our design of an efficient tutorial for increasing ASL-SRT raters' metalinguistic skills in detecting and categorizing response behavior.

The ASL-SRT holds promise as a research tool for the investigation of sign language processing across a variety of populations. In addition, the test can be applied to the screening, detection, and diagnosis of language behavior related to second language learning, language transfer and L1 intrusion, and age of acquisition issues, as well as for the detection and diagnosis of language impairment among native signers. Our future plans include presenting the ASL-SRT test to additional deaf native signers of varying ages, L2 hearing signers, late-learning congenitally deaf signers, and late-deafened signers as they progress through different levels of fluency in learning ASL. Such data will provide

additional information on the heuristics used at different levels of fluency and knowledge of signed languages.

REFERENCES

Baddeley, A. D. (1995). "Working memory," in *The Cognitive Neurosciences*, ed M. S. Gazzaniga (Cambridge: MIT Press), 755–764.

Baddeley, A. D., and Hitch, G. J. (2007). "Working Memory: past, present and future?," in *The Cognitive Neuroscience of Working Memory,* eds N. Osaka, R. H. Logie, and M. D'esposito (New York, NY: Oxford University Press), 1–20.

Bavelier, D., Newport, E. L., Hall, M. L., Supalla, T., and Boutla, M. (2006). Persistent differences in short-term memory span between sign and speech: implications for cross-linguistic comparisons. *Psychol. Sci.* 17, 1090–1092. doi: 10.1111/j.1467-9280.2006.01831.x

Boutla, M., Supalla, T., Newport, E. L., and Bavelier, D. (2004). Short-term memory span: insights from sign language. *Nat. Neurosci.* 7, 997–1002. doi: 10.1038/nn1298

Brentari, D. (1998). *A Prosodic Model of Sign Language Phonology.* Cambridge, MA: MIT Press.

Fischer, S. D., and Janis, W. (1990). "Verb sandwiches in ASL," in *Current Trends in European Sign Language Research*, eds.S. Prillwitz and T. Vollhaber (Hamburg: Signum Press), 279–294.

Fodor, J. A. (1983). *The Modularity of Mind.* Cambridge, MA: Bradford.

Haarmann, H. J., Davelaar, E. J., and Usher, M. (2003). Individual differences in semantic short-term memory capacity and reading comprehension. *J. Mem. Lang.* 48, 320–345. doi: 10.1016/S0749-596X(02)00506-5

Hammill, D. D., Brown, V. L., Larsen, S. C., and Wiederholt, J. L. (1994). *Test of Adolescent and Adult Language, 3rd Edn.* Austin, TX: PRO-ED, Inc.

Hauser, P. C., Paludneviciene, R., Supalla, T., and Bavelier, D. (2008). "American Sign Language-Sentence reproduction test: development and implications," in *Sign Language: Spinning and Unraveling the Past, Present and Future*, ed R. M. de Quadros (Petropolis: Editora Arara Azul), 160–172.

Just, M. A., and Carpenter, P. A. (1992). A capacity theory of comprehension: individual differences in working memory. *Psychol. Rev.* 99, 122–149. doi: 10.1037/0033-295X.99.1.122

Lewis, R. L., Vasishth, S., and Van, D. J. (2006). Computational principles of working memory in sentence comprehension. *Trends Cogn. Sci.* 10, 447–454. doi: 10.1016/j.tics.2006.08.007

Liddell, S., and Johnson, R. (1989). American Sign Language: the phonological base. *Sign Lang. Stud.* 64, 197–277.

Mayberry, R. (2010). "Early language acquisition and adult language ability: what sign language reveals about the critical period for language," in *The Oxford Handbook of Deaf Studies, Language, and Education*, Vol. 2, eds M. Marschark and P. Spencer (Oxford: Oxford University Press), 281–291.

Mayberry, R., and Fischer, S. (1989). Looking through phonological shape to lexical meaning: the bottleneck of non-native sign language processing. *Mem. Cogn.* 17, 740–754. doi: 10.3758/BF03202635

McElree, B., Foraker, S., and Dyers, L. (2003). Memory structures that subserve sentence comprehension. *J. Mem. Lang.* 48, 67–91. doi: 10.1016/S0749-596X(02)00515-6

Morford, J. P. (2002). "Why does exposure to language matter?," in *The Evolution of Language Out of Pre-Language*, eds T. Givon and B. Malle (Philadelphia, PA: John Benjamins Publishing Company), 329–341.

Newport, E. L. (1990). Maturational constraints on language learning. *Cogn. Sci.* 14, 11–28. doi: 10.1207/s15516709cog1401_2

Newport, E. L., and Meier, R. (1985). "The acquisition of American Sign Language," in *The Crosslinguistic Study of Language Acquisition*, Vol. 1, The Data, ed Dan I. Slobin (Hillsdale, NJ: Lawrence Erlbaum Associates), 881–938.

Potter, M. C. (1990). "Remembering," in *An invitation to Cognitive Science,* Vol. 3, eds D. Osherson and E. Smith (Cambridge, MA: Bradford/MIT Press), 3–32.

Potter, M. C., and Lombardi, L. (1990). Regeneration in the short-term recall of sentences. *J. Mem. Lang.* 29, 633–654. doi: 10.1016/0749-596X(90)90042-X

Ronnberg, J., Rudner, M., Foo, C., and Lunner, T. (2008). Cognition counts: a working memory system for ease of language understanding (ELU). *Int. J. Audiol.* 47(Suppl. 2), S171–S177. doi: 10.1080/14992020802301167

Sandler, W., and Lillo-Martin, D. (2006). *Sign Languages and Universals*. Cambridge: Cambridge University Press.

Supalla, T. (1990). "Serial verbs of motion in ASL," in *Theoretical Issues in Sign Language Research*, Vol. 1, eds S. D. Fischer and P. Siple (Chicago: The University of Chicago Press), 127–152.

How sensory-motor systems impact the neural organization for language: direct contrasts between spoken and signed language

Karen Emmorey[1]*, Stephen McCullough[1], Sonya Mehta[2,3] and Thomas J. Grabowski[3]

[1] *Laboratory for Language and Cognitive Neuroscience, School of Speech, Language, and Hearing Sciences, San Diego State University, San Diego, CA, USA*
[2] *Department of Psychology, University of Washington, Seattle, WA, USA*
[3] *Department of Radiology, University of Washington, Seattle, WA, USA*

To investigate the impact of sensory-motor systems on the neural organization for language, we conducted an $H_2{}^{15}O$-PET study of sign and spoken word production (picture-naming) and an fMRI study of sign and audio-visual spoken language comprehension (detection of a semantically anomalous sentence) with hearing bilinguals who are native users of American Sign Language (ASL) and English. Directly contrasting speech and sign production revealed greater activation in bilateral parietal cortex for signing, while speaking resulted in greater activation in bilateral superior temporal cortex (STC) and right frontal cortex, likely reflecting auditory feedback control. Surprisingly, the language production contrast revealed a relative increase in activation in bilateral occipital cortex for speaking. We speculate that greater activation in visual cortex for speaking may actually reflect cortical attenuation when signing, which functions to distinguish self-produced from externally generated visual input. Directly contrasting speech and sign comprehension revealed greater activation in bilateral STC for speech and greater activation in bilateral occipital-temporal cortex for sign. Sign comprehension, like sign production, engaged bilateral parietal cortex to a greater extent than spoken language. We hypothesize that posterior parietal activation in part reflects processing related to spatial classifier constructions in ASL and that anterior parietal activation may reflect covert imitation that functions as a predictive model during sign comprehension. The conjunction analysis for comprehension revealed that both speech and sign bilaterally engaged the inferior frontal gyrus (with more extensive activation on the left) and the superior temporal sulcus, suggesting an invariant bilateral perisylvian language system. We conclude that surface level differences between sign and spoken languages should not be dismissed and are critical for understanding the neurobiology of language.

Keywords: American Sign Language, audio-visual English, bimodal bilinguals, PET, fMRI

INTRODUCTION

Evidence from lesion-based, neuroimaging, and neurophysiological studies has revealed that the same left perisylvian regions are recruited during the production and comprehension of both spoken and signed languages (for reviews see Emmorey, 2002; MacSweeney et al., 2008a; Corina et al., 2013). Nonetheless, the neural substrates for speech and sign are not identical. In the experiments presented here, we endeavor to identify the specific sensory- and motor-related systems that are differentially recruited for spoken and signed languages within the same individual: hearing bilinguals who are native users of American Sign Language (ASL) and English. We first examine language production using positron emission tomography (PET) and report the first study (to our knowledge) to contrast sign and word production within participant and without reference to a common motoric baseline that would remove the modality effects of interest. We next review previous production studies that identified the neural overlap between signing and speaking in order to provide

a complete picture of language produced by hand and by mouth. We then turn to language comprehension and report the results of an fMRI study that directly contrasts sentence comprehension in ASL and English in hearing bilinguals. Finally, we present data from this study that reveals for the first time the neural conjunction for visual sign comprehension and audiovisual speech comprehension.

The goal of these direct contrasts for both language production and comprehension is to target the neural substrates that are specific to visual-manual and auditory-vocal languages. The goal of the conjunction analyses is to identify neural substrates that are common to both language types. By establishing both the differences and similarities between the neural substrates that support spoken and signed language processing, we can characterize the neurobiological impact of using the hands or the vocal tract as the primary linguistic articulators and of the perceptual reliance on vision or audition for language comprehension.

EXPERIMENT 1: CONTRASTING THE NEURAL SUBSTRATE FOR SIGN vs. WORD PRODUCTION

To date, no neuroimaging study has directly contrasted signing and speaking without subtracting activation from a common motor baseline. For example, Emmorey et al. (2007) conducted a between-group comparison of deaf signers and hearing speakers in which participants overtly named pictures in contrast to a baseline task that required a manual or a vocal response. The goal of that group comparison was to investigate similarities and differences between sign and word production at higher levels of lexical processing, using the baseline task in part to eliminate surface-level differences between sign and speech articulation. In fact, Emmorey et al. (2007) reported no neural regions that exhibited greater activity for speech compared to sign when controlling for low-level motoric and sensory differences through the use of baseline tasks.

Braun et al. (2001) contrasted signed and spoken narrative production (spontaneous autobiographical narratives) by hearing ASL-English bilinguals. Like the Emmorey et al. (2007) study, the contrast between the production of English and ASL was conducted with respect to perceptual-motoric baseline tasks for speech (oral movements with vocalizations) and for sign (hand and limb movements). Signing and speaking were not directly contrasted with one another, although the interaction analyses suggested activations that Braun et al. (2001) attributed to modality-dependent features related to articulation. Specifically, for English increased neural activity was observed in prefrontal and subcortical regions, and Braun et al. (2001) hypothesized that greater activity in these regions reflected the more rapid, sequential oral articulations required for speech. ASL production was associated with greater activity in the superior parietal and paracentral lobules, which Braun et al. (2001) attributed to the execution of complex handshapes and movements to various locations on the body. However, Braun et al. (2001) acknowledged that some of these differences might also reflect modality-specific differences at higher levels of processing, such as the syntactic and semantic use of signing space.

By directly contrasting speaking and signing, we can identify what perceptual and articulatory differences are found when the sensory-motor activation related to vocal and manual baselines is not "subtracted out" of the analysis; that is, a direct contrast provides a better assessment of the neural differences that are specifically related to the sensory-motoric properties of sign vs. word production. Further, we can determine whether the sensory-motoric differences identified by Braun et al. (2001) during narrative production also occur during the production of single lexical items.

MATERIALS AND METHODS
Participants
Fourteen ASL-English bilinguals participated in an $H_2{}^{15}O$-PET study (9 women; mean age = 27 years). All participants were exposed to ASL from birth from their deaf signing families, reported normal hearing, were right-handed, and had 12 or more years of formal education.

Materials and procedure
Participants overtly named pictures (line drawings of objects from Bates et al., 2003) in either ASL or in English. For each language, participants named 80 pictures in four blocks of 20 pictures each; half had high and half had low frequency names[1]. The order of the ASL and English naming conditions was counterbalanced across participants, and each picture was only presented once during the experiment (i.e., half of the participants named a given picture in English and half in ASL). Pictures were presented to participants using I-glasses SVGA Pro goggles (I-O Display Systems; Sacramento, CA). For each naming block, the picture stimuli were presented from 5 s after the injection (approximately 7–10 s before the bolus arrived in the brain) until 35 s after the intravenous bolus injection of 15 mCi of [15O]water, and each picture was presented for 1 s followed by a 1 s inter-stimulus-interval.

Image acquisition
All participants underwent MR scanning in a 3.0T TIM Trio Siemens scanner to obtain a 3D T1-weighted structural scan with isotropic 1 mm resolution using the following protocol: MP-RAGE, TR 2530, TE 3.09, TI 800, FOV 25.6 cm, matrix $256 \times 256 \times 208$. PET data were acquired with a Siemens/CTI HR+ PET system using the following protocol: 3D, 63 image planes, 15 cm axial FOV, 4.5 mm transaxial and 4.2 mm axial FWHM resolution.

Images of rCBF were computed using the [15O]water autoradiographic method (Herscovitch et al., 1983; Hichwa et al., 1995) as follows. Dynamic scans were initiated with each injection and continued for 100 s, during which 20 5-s frames were acquired. To determine the time course of bolus transit from the cerebral arteries, time-activity curves were generated for regions of interest placed over major vessels at the base of the brain. The eight frames representing the first 40 s immediately after transit of the bolus from the arterial pool were summed to make an integrated 40-s count image. These summed images were reconstructed into 2 mm pixels in a 128×128 matrix.

Spatial normalization
PET data were spatially normalized to a Talairach-compatible atlas through a series of coregistration steps (see Damasio et al., 1994; Grabowski et al., 1995, for details). Prior to registration, the MR data were manually traced to remove extracerebral voxels. Talairach space was constructed directly for each participant via user-identification of the anterior and posterior commissures and the midsagittal plane on the 3D MRI data set in Brainvox. An automated planar search routine defined the bounding box and piecewise linear transformation was used (Frank et al., 1997), as defined in the Talairach atlas. After Talairach transformation, the MR data sets were warped (AIR 5th order non-linear algorithm) to an atlas space constructed by averaging 50 normal

---

[1] The frequency manipulation generated only weak differences in neural activity, and the results reported here are collapsed across frequency. Word length and sign length were also manipulated (one syllable vs. two syllables), but no significant length effects were observed for either language, and the reported results are collapsed across word/sign length.

Talairach-transformed brains, rewarping each brain to the average, and finally averaging them again, analogous to the procedure described in Woods et al. (1999). Additionally, the MR images were segmented using a validated tissue segmentation algorithm (Grabowski et al., 2000), and the gray matter partition images were smoothed with a 10 mm FWHM Gaussian kernel. These smoothed gray matter images served as the target for registering participants' PET data to their MR images.

For each participant, PET data from each injection were coregistered to each other using Automated Image Registration (AIR 5.25, Roger Woods, UCLA). The coregistered PET data were averaged, and the mean PET image was then registered to the smoothed gray matter partition using FSL (Jenkinson and Smith, 2001; Jenkinson et al., 2002). The deformation fields computed for the MR images were then applied to the PET data to bring them into register with the Talairach-compatible atlas. After spatial normalization, the PET data were smoothed with a $16.1 \times 16.1 \times 15.0$ mm FWHM Gaussian kernel using complex multiplication in the frequency domain to produce a final isotropic voxel resolution of 18 mm. PET data from each injection were normalized to a global mean of 1000 counts per voxel.

Regression analysis

PET data were analyzed with a pixelwise general linear model (Friston et al., 1995). Regression analysis was performed using tal_regress, a customized software module based on Gentleman's least squares routines (Miller, 1991) and cross-validated against SAS (Grabowski et al., 1996). The regression model included covariables for task condition (language modality, frequency, and length manipulations) and subject effects. The contrast between signing and speaking was computed using the appropriate linear combination of task conditions. Results were thresholded for a two tailed t-test (familywise error rate $p < 0.05$) using random field theory (RFT) to correct for multiple spatial comparisons across the whole brain (Worsley et al., 1992; Worsley, 1994).

RESULTS

Table 1 provides the local maxima for the direct contrast between sign production and word production, and these results are illustrated in Figure 1. As expected from previous studies, sign production was associated with greater activation in parietal cortices compared to speaking, while speaking resulted in greater activation in bilateral superior temporal cortices, which is most likely due to the auditory feedback that occurs during speaking. In addition, differences within sensory-motor cortices were observed reflecting articulatory differences between signing and speaking. For signing, there was greater activation bilaterally in the cerebellum and in superior regions of the pre- and post-central gyri associated with motor and somatosensory responses for the upper extremities of both limbs. For speaking, there was increased activation in more inferior sensory-motor regions associated with control of the face and mouth. Spoken word production also resulted in increased activation in bilateral middle and superior frontal cortices, compared to sign production.
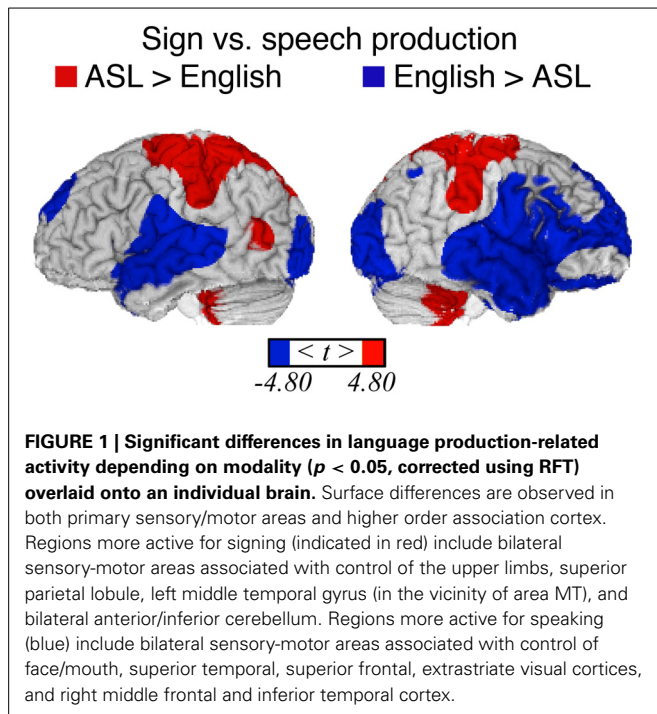
Somewhat surprisingly, more extensive activation in bilateral occipital cortex was observed for speaking in contrast to signing. To confirm this unexpected result, we conducted a conjunction

Table 1 | Summary of PET activation results for the comparison between signing and speaking.

Region	Side	X	Y	Z	T
SIGNING > SPEAKING					
Frontal cortex					
Pre-central gyrus (BA 6)	R	+27	−17	+59	9.87
Temporal cortex					
Mid. temporal gyrus (BA 21)	L	−43	−62	+11	11.99
Parietal cortex					
Inferior parietal cortex (BA 40)	R	+56	−29	+38	10.24
Superior parietal lobule (BA 1, 2, 3, 4, 7)	L	−35	−31	+50	22.25
	R	+33	−35	+51	11.03
Occipital cortex					
Cuneus (BA 19)	L	−10	−81	+36	5.76
Subcortical regions					
Thalamus	L	−9	−19	+2	5.55
Cerebellum					
	L	−33	−39	−28	9.67
	R	+18	−45	−19	11.70
SPEAKING > SIGNING					
Frontal cortex					
Pre-central gyrus (BA 6)	L	−56	−6	+42	−6.16
	R	+58	−2	+40	−8.32
Inferior frontal gyrus (BA 47)	R	+29	+21	−17	−7.51
Middle frontal gyrus (BA 8)	R	+52	+16	+38	−6.16
Middle frontal gyrus (BA 9)	L	−54	+18	+30	−5.12
Inf./Mid. frontal gyrus (BA 46)	L	−52	+36	+8	−4.82
	R	+55	+25	+26	−8.24
Medial frontal gyrus (BA 10)	R	+16	+54	+14	−8.87
Medial frontal gyrus (BA 11)	L	−8	+49	−14	−7.02
Superior frontal gyrus (BA 9)	R	+5	+47	+32	−10.18
Superior frontal gyrus (BA 11)	R	+16	+45	−14	−6.41
Temporal cortex					
Superior temporal gyrus	L	−58	−16	+5	−18.89
	R	+63	−10	+7	−18.10
Inferior parietal lobule (BA 40)	R	+47	−60	+48	−5.33
Occipital cortex					
Mid. occipital gyrus (BA 18)	R	+32	−93	+15	−10.39
Inf. occipital gyrus (BA 18)	L	−29	−96	−6	−10.57
Lingual gyrus (BA 18)	R	+24	−98	−6	−9.82

Results are from the whole brain analysis [critical $t_{(91)} = \pm 4.80$].

analysis using the data from Emmorey et al. (2005) in which a different group of hearing bilinguals named pictures in either ASL or English. In that study, bilinguals viewed line drawings depicting a spatial relation between two objects and produced either an ASL locative classifier construction or an English preposition that described the spatial relation, and the comparison task was to name the figure object (colored red) in either ASL or in English. No motoric baseline was included in this study, and Emmorey et al. (2005) did not report a direct contrast between sign and speech because their focus was the neural correlates of spatial language in ASL compared to English. To compute the contrast between signing and speaking, PET data from the object-naming condition in the Emmorey et al. (2005) study were processed in

Sign vs. speech production
■ ASL > English   ■ English > ASL

< t >
-4.80   4.80

FIGURE 1 | Significant differences in language production-related activity depending on modality ($p < 0.05$, corrected using RFT) overlaid onto an individual brain. Surface differences are observed in both primary sensory/motor areas and higher order association cortex. Regions more active for signing (indicated in red) include bilateral sensory-motor areas associated with control of the upper limbs, superior parietal lobule, left middle temporal gyrus (in the vicinity of area MT), and bilateral anterior/inferior cerebellum. Regions more active for speaking (blue) include bilateral sensory-motor areas associated with control of face/mouth, superior temporal, superior frontal, extrastriate visual cortices, and right middle frontal and inferior temporal cortex.

an essentially identical manner as the current data. Results were thresholded for a two tailed t-test (familywise error rate $p < 0.05$, corrected using RFT; Worsley et al., 1992; Worsley, 1994). We used the Minimum Statistic compared to the Conjunction Null method, as described in Nichols et al. (2005) because this type of conjunction analysis is by nature conservative, requiring identified regions to be independently significant in both groups of subjects. This conjunction analysis replicated and confirmed the surprising finding that when directly contrasted, greater activation in bilateral occipital cortex was observed for speaking than for signing (see Supplementary Table).

DISCUSSION

Differences between the linguistic articulators for speaking and signing were reflected in greater activation along inferior regions of the sensory-motor strip associated with the oral articulators for speech and increased activation in superior regions associated with the arms for sign production. We did not see evidence for greater engagement of the prefrontal corticostriatal-thalamocortical circuit for speech that Braun et al. (2001) hypothesized to be preferentially recruited to control the timing and sequencing of phonetic units when speaking. However, the timing demands for speaking are likely to be greater for connected narratives than for the production of isolated individual words.

Spoken word production results in auditory feedback, which is reflected in more activation within bilateral superior temporal cortex (STC) for speaking. In addition, spoken word production recruited right frontal cortices to a greater extent than sign production (see **Figure 1**; **Table 1**). Listening to speech, including self-produced speech, activates right inferior frontal cortex (see **Figure 3** below and Tourville et al., 2008), whereas self-produced signing does not result in a visual signal that is parallel to

perceiving sign language produced by another person (Emmorey et al., 2009a,b), and self-produced signing does not strongly activate right frontal cortices (e.g., Emmorey et al., 2003; Hu et al., 2011). The activation peak in the right middle frontal gyrus ($+52$, $+16$, $+38$) for speaking ($>$signing) is near the coordinates for the right lateralized feedback control component for speech production proposed by the DIVA model (Tourville and Guenther, 2011). According to this model, right ventral premotor and right inferior frontal cortex (pars triangularis) receive auditory feedback signals from left and right posterior superior temporal gyri. These right frontal regions mediate between auditory and motor cortices during self-monitoring of speech production. It is unlikely that self-monitoring of sign production relies on this feedback circuit; rather, sign monitoring appears to be more dependent on somatosensory than visual feedback (Emmorey et al., 2009a,b), which likely relies on a fronto-parietal-cerebellar circuit.

The direct contrast between speaking and signing revealed a surprising relative increase in activation within bilateral occipital cortex for speaking. We speculate that greater activation in visual cortex for speaking in contrast to signing may reflect the suppression of activation in these areas when signing. That is, the neural response to self-produced signing within visual cortex may be suppressed, just as the neural response in auditory cortex is suppressed during self-produced speech (e.g., Numminen et al., 1999; Houde et al., 2002). Note that Braun et al. (2001) required participants to close their eyes when speaking and signing, and thus this study would be unable to detect modulations in occipital cortex arising from visual input during language production.

Neural responses to visual input may be generally attenuated during signing in order to help distinguish self-generated motion toward the body from "externally generated" movements of hands or arms toward the body. A signer (or speaker) may be more likely to flinch when another person's hand moves rapidly toward the face or body than when such hand movement is self-produced. Similarly, Hesse et al. (2010) reported cortical attenuation of somatosensory activation elicited by self-produced tactile stimulation and argued that motor commands generate sensory expectations that are compared with the actual sensory feedback to allow for the distinction between internally and externally generated actions. It is possible that posterior parietal cortex and/or left MT (regions that are more active during signing than speaking) may actively inhibit the neural response in occipital cortex to self-generated hand and arm movements during signing. Such modulation could reduce visual attention to self-generated hand movements during signing, and such modulation of occipital cortex would not occur during speaking. However, further investigation is needed to support this speculative hypothesis.

Consistent with several other studies (Braun et al., 2001; Corina and Knapp, 2006; Emmorey et al., 2007), sign production resulted in greater activation in parietal cortex, with more extensive activation in the left hemisphere. The probable source of activation in anterior parietal cortex (including the post-central gyrus) is the somatosensory and proprioceptive feedback received during sign production. Posterior parietal cortex is engaged during the voluntary production of motor movements of the hand and arm, including reaching, grasping, and tool-use

(see Creem-Regehr, 2009, for review). Phonological encoding for sign language requires the selection and assembly of one to two hand configurations, locations on the face or body, and movement trajectories. Although inferior parietal cortex is involved in sensory-motor integration during speech production (e.g., Hickok et al., 2009), inferior parietal cortex may play a greater role in sign than speech production. Furthermore, the direct contrast reported here indicates that right inferior parietal cortex is relatively more engaged in sign production (see **Figure 1**).

Activation in the anterior cerebellum was also greater for signing than speaking, and this region is thought to be involved in sensorimotor processing and prediction of hand and arm actions (e.g., Lorey et al., 2010). Greater cerebellar activity for signing likely reflects the greater demands of on-line motor control for the fingers, hands, and arms. This result is also consistent with recent evidence from diffusion tensor imaging indicating higher fractional anisotropy in the cerebellum for deaf signers relative to hearing non-signers (Tungaraza et al., 2011).

As expected based on previous (non-direct) comparisons between speaking and signing, there was no significant difference between sign and word production in the left inferior frontal gyrus [Brodmann area (BA) 44/45]. There was also no significant difference between language modalities within left posterior temporal cortex, with the exception that sign production engaged left MT to a greater extent than speaking (see **Figure 1**). Activation in left MT might reflect linguistic processing of hand movements seen in peripheral vision during sign production and/or involvement in phonological encoding of movement parameters for signs. Several studies have found a strong left hemisphere asymmetry for motion processing for signers (both deaf and hearing) compared to non-signers (e.g., Bavelier et al., 2001; Bosworth and Dobkins, 2001).

In sum, when vocal and manual baselines are not included in the direct contrast between speaking and signing, clear modality-related differences in cortical activation emerge. Auditory feedback during speech production engaged STC bilaterally, as well as right inferior frontal cortex. Sign production engaged parietal and cerebellar cortices to a greater extent than speaking, reflecting neural control required to articulate target hand configurations and produce directed movements of the hand and arm toward the body and in space. In addition, the direct contrast between speaking and signing revealed a surprising relative increase in activation within bilateral occipital cortex for speaking, which was confirmed through a conjunction analysis using results from a separate group of ASL-English bilinguals (see Supplementary Table). We speculate that this finding actually reflects the suppression of activation in visual areas when signing, just as neural responses in auditory cortex are suppressed during self-produced speech. Finally, it is worth noting that for spoken language ("unimodal") bilinguals, the production of their two languages relies on essentially the same neural substrate with few differences, particularly for early simultaneous bilinguals (e.g., Simmons et al., 2011; Parker Jones et al., 2012). The direct contrast between signing and speaking shown in **Figure 1** illustrates the rather dramatic difference in neural resources required for the production of a bimodal bilingual's two languages (see Emmorey and McCullough, 2009, for further discussion of the neural consequences of bimodal

bilingualism). We now turn to the similarities between language produced by mouth and by hand.

COMMON NEURAL SUBSTRATES FOR SIGN AND WORD PRODUCTION

The design of our PET studies with hearing ASL-English bilinguals did not permit a conjunction analysis for speech and sign production because no sensory-motoric or fixation baselines were included (conjunction analyses require a reference baseline). The original questions addressed by our studies required only within condition contrasts between lexical types (e.g., high vs. low frequency items or prepositions vs. nouns), and thus we opted not to include additional injections for a baseline condition. However, other studies have specifically identified the neural overlap for signing and speaking using baseline measures, and we briefly summarize those results.

Braun et al. (2001) asked bimodal bilinguals to produce autobiographical narratives in either English or ASL and to perform non-meaningful complex and simple oral-facial or manual-brachial movements as baseline controls while undergoing PET imaging. Conjunction analyses revealed that discourse production for both languages relied on classical left perisylvian language regions: inferior frontal cortex and posterior STC, extending into middle temporal gyrus. Shared activation for sign and speech production also extended beyond these classical language regions, including left anterior insula, right posterior superior temporal gyrus (STG) extending into the angular gyrus, and bilateral basal temporal cortex (fusiform and lingual gyri).

Braun et al. (2001) suggest that left anterior brain regions [the frontal operculum, insula, lateral premotor cortex, and supplementary motor area (primarily pre-SMA)] are involved the phonological and phonetic encoding of complex articulatory movements for both speaking and signing. These same regions were also reliably activated by the complex oral and limb motor tasks, suggesting that language formulation was not required to engage these anterior brain regions. Of course, this finding does not imply that these anterior cortical regions only play a motor-articulatory role in language production—rather, they point to their multifunctionality, particularly the frontal operculum (cf. Grodzinsky and Amunts, 2005). In contrast, bilateral posterior brain regions (posterior superior and middle temporal gyri, posterior superior temporal sulcus, and angular gyrus) were only engaged during language production and not during complex motor baseline tasks. Braun et al. (2001) suggest that these bilateral posterior brain regions are involved in semantic and pragmatic processes required to create autobiographical narratives in both ASL and English.

Emmorey et al. (2007) conducted a conjunction analysis for single sign production (by native deaf ASL signers) and single word production (by hearing English speakers) in a picture-naming task, with a baseline task that required participants to make an orientation judgment (upright or inverted) for unknown faces, overtly signing or saying *yes* or *no* on each trial. Consistent with the Braun et al. (2001) results, both sign and speech engaged the left inferior frontal gyrus (Broca's area) indicating a

modality-independent role for this region in lexical production. Using probabilistic cytoarchitectonic mapping and data from the Braun et al. (2001) study, Horwitz et al. (2003) reported that BA 45 was engaged during both speaking and signing, but there was no involvement of BA 44, compared to the motor baseline conditions. In addition, there was extensive activation in BA 44, but not in BA 45, for the non-linguistic oral and manual control tasks compared to rest. This pattern of results suggests that BA 44, rather than BA 45, is engaged during the production of complex movements of the oral and manual articulators and that BA 45 is more likely engaged in articulator-independent aspects of language production. Finally, Emmorey et al. (2007) found that both sign and word production engaged left inferior temporal regions, which have been shown to be involved in conceptually driven lexical access (e.g., Indefrey and Levelt, 2004).

Overall, these conjunction studies, along with additional data from lesion and neuroimaging studies, indicate that sign and speech production both rely on a primarily left lateralized neural network that includes left inferior frontal cortex (BA 44/45, 46, and 47), pre-SMA, insula, middle/inferior temporal cortex, and inferior parietal cortex (see also Hickok et al., 1996; Corina et al., 2003; Kassubek et al., 2004). We point out that our null findings for the direct contrast between speech and signing in this left lateralized network are consistent with the conjunction study results.

EXPERIMENT 2: CONTRASTING THE NEURAL SUBSTRATE FOR SIGNED vs. SPOKEN LANGUAGE COMPREHENSION

As with language production, few studies have directly contrasted signed and spoken language comprehension. An early PET study by Söderfelt et al. (1997) presented hearing bilinguals with short, signed narratives (Swedish Sign Language) and audiovisually presented spoken Swedish narratives (a video of the same model speaking). The direct contrast revealed greater activation in bilateral perisylvian cortex for audiovisual speech comprehension and greater activation in bilateral middle/inferior temporal cortex (BA 37, 19) for sign language comprehension, reflecting auditory neural responses for speech perception and visual motion processing for sign perception. No other differences were reported, but this study was underpowered with only six participants and without the spatial resolution and sensitivity of modern fMRI. For example, it is possible that parietal cortex may have been more involved in signed than spoken language comprehension given the role of parietal cortex in sign production and in the recognition of human actions (e.g., Corina and Knapp, 2006), but the Söderfelt et al. (1997) study may have been unable to detect this difference. Neuroimaging studies that have separately examined sign language comprehension (by deaf or hearing signers) and audiovisual spoken language comprehension have observed more parietal activation for sign comprehension (e.g., MacSweeney et al., 2002a). Here we report the first direct contrast (to our knowledge) between the comprehension of sign language and audiovisual spoken language by hearing native ASL-English bilinguals. We also report the first conjunction analysis (to our knowledge) that identifies the neural overlap between the two languages for these bilinguals.

MATERIALS AND METHODS
Participants
Thirteen hearing native ASL-English bilinguals (9 females; mean age = 26.4 years; SD = 4.7 years) participated in the study. All participants were born into deaf signing families, were right handed, and had normal or corrected-to-normal vision by self-report. ASL data from these participants was presented in McCullough et al. (2012).

Materials and procedure
The spoken language materials are from Saygin et al. (2010) and consisted of audiovisual English sentences produced by a female native speaker that expressed motion (e.g., "The deer jumped over the brook"), static location (e.g., "Her family lives close to the river"), or metaphorical (fictive) motion (e.g., "The hiking trail crossed the barren field"). Co-speech gestures were not produced. The signed language materials are from McCullough et al. (2012) and consisted of similar (but non-identical) ASL sentences produced by a male native signer that expressed motion (e.g., English translation: "Many dogs were running loose around the farmyard") or static location (e.g., English translation: "The lion slept in his enclosure at the zoo"). For the purposes of this study, sentence type was not a treated as a variable of interest.

Presentation of English and ASL sentences was counterbalanced across participants. Participants pressed a button when they heard/saw a sentence that was semantically anomalous (e.g., "The wooden fence crosses the late curfew."). Anomalous sentences were relatively rare, occurring either once or never within a block, and frequency was matched across languages (12% of sentences were anomalous for both ASL and English). The baseline condition for ASL consisted of video clips of the model signer sitting in the same position but not signing, and participants decided whether the color of a black dot superimposed on the model's chin changed to white during the baseline. The baseline condition for English was parallel: participants saw video clips of the same speaker sitting in the same position, but remaining silent and with a dot superimposed on her chin. Participants monitored whether a continuous pure tone presented along with the video stimuli changed frequency, and the change in frequency occurred simultaneously with the change in dot color. The (in)frequency of the dot targets was matched with the sentence condition targets (12%). These low-level baseline conditions presented visual (and auditory) stimuli along with a simple button press task to provide a reference against which to measure neural responses to the English and ASL sentences.

MRI data acquisition and analysis
MRI data were collected using a 3-Tesla GE Signa Excite scanner equipped with an eight-element phased-array head coil at the Center for fMRI at the University of California, San Diego. For each participant, a 1 × 1 × 1.3 mm anatomical scan was collected, usually in the middle of the scanning session. Echo-planar volumes were acquired from the whole brain with a repetition time (TR) of 2000 ms, an echo time (TE) of 30 ms, 3.5 mm in-plane resolution, and 4 mm slice thickness (no gap). Image preprocessing and statistical analyses were performed using Analysis of Functional Neuroimages (AFNI) software package (version

AFNI_2010_10_19_1028; Cox, 1996). Further details on data acquisition and pre-processing can found in Saygin et al. (2010) and McCullough et al. (2012).

For the individual-level analysis, ASL and English sentence blocks were modeled as regressors of interest in the design matrix with respect to the control baseline. The design matrix was constructed using AFNI's 3dDeconvolve. Six motion parameters, obtained during head motion correction (AFNI's 3dvolreg), and a Legendre polynomial set ranging from zero to third order to account for slow drifts were included in the design matrix as nuisance regressors. The regressor of interest beta values and t-values from each individual were estimated and calculated using AFNI's 3dREMLFIT (Chen et al., 2012). For the group-level analysis, individuals' voxelwise betas and their corresponding t-values for each contrast of interest served as inputs to group-level, mixed-effects meta-analysis (AFNI's 3DMEMA, Chen et al., 2012). We used false discovery rate correction for multiple comparisons to identify clusters of significant activation in the ASL vs. English sentence contrast. Only clusters of 30 or more contiguous voxels surviving $q = 0.001$ are reported.

To identify the regions of the common activation between ASL and English sentence comprehension relative to the baseline, a conjunction analysis was performed using the minimum statistic ($q = 0.01$) for each condition to test the conjunction null hypothesis (i.e., minimum statistic compared to conjunction null; Nichols et al., 2005).
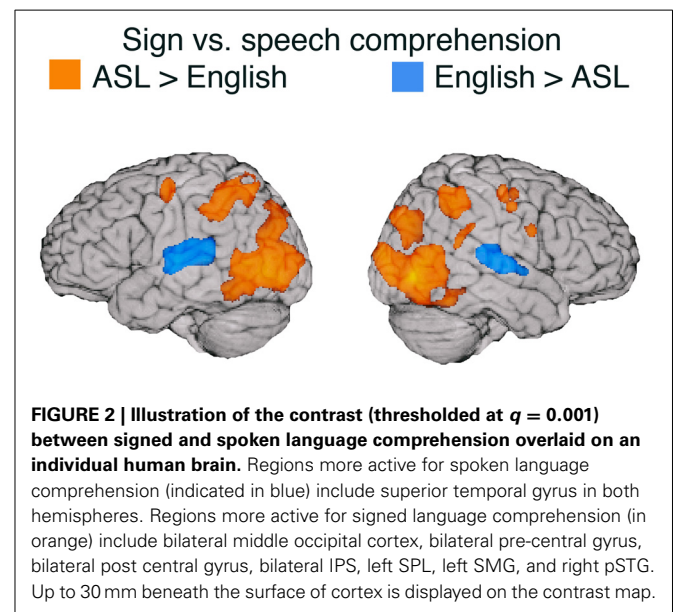
RESULTS

Table 2 lists the peak Talairach coordinates and cluster volumes for the contrast between ASL and English, and the results are illustrated in **Figure 2**. Only the STG (bilaterally) was more active for comprehension of spoken than signed language. In contrast, several regions were more active for the comprehension of signed than spoken language: bilateral posterior middle temporal cortex (extending into lateral occipital cortex), bilateral inferior and superior parietal cortices (more extensive on the left), and bilateral premotor cortex.

DISCUSSION

Replicating Söderfelt et al. (1997), the audio-visual signal for speech activated the STG bilaterally to a greater extent than the purely visual signal for sign language for hearing ASL-English bilinguals (see **Figure 2**). Although there is evidence that visual stimuli (including sign language) activate auditory cortex for deaf people (e.g., Finney et al., 2001; Cardin et al., 2013), comprehending spoken language for hearing individuals requires significantly more neural resources and sustained activation in auditory cortices compared to comprehending sign language (see also Leonard et al., 2012). In addition, MacSweeney et al. (2002a) found that hearing native users of British Sign Language (BSL) exhibited less extensive activation along left STG compared to deaf native signers when comprehending BSL sentences and hypothesized that auditory processing of speech has privileged access to more anterior regions of STG (adjacent to primary auditory cortex), such that hearing signers engage this region much less strongly during sign language processing (see also Emmorey and McCullough, 2009).

Table 2 | Peak Talairach coordinates and cluster volumes for the contrast between sign and spoken language comprehension ($q = 0.001$).

Region	Side	X	Y	Z	Vol mm³	T
SIGN > SPOKEN LANGUAGE						
Frontal cortex						
Pre-central gyrus (BA 4)	L	−23	−11	+50	964	6.09
	R	+50	+2	+26	281	6.14
	R	+34	−12	+46	972	5.93
Temporal cortex						
Posterior superior temporal gyrus (BA 22)	R	+57	−40	+23	589	6.26
Parietal cortex						
Post central sulcus (BA 3)	L	−27	−46	+44	4161	6.96
	R	+30	−46	+44	3648	7.83
Intraparietal sulcus (BA 39)	L	−24	−76	+29	2221	6.75
	R	+28	−76	+30	2804	8.53
Superior parietal lobule (BA 7)	L	−19	−61	+48	1188	6.73
Supramarginal gyrus (BA 40)	L	−53	−35	+37	527	5.77
Occipital cortex						
Middle occipital gyrus (BA 19)	L	−36	−70	+4	17172	13.09
	R	+41	−69	+1	17120	24.76
SPOKEN > SIGN LANGUAGE						
Temporal cortex						
Superior temporal gyrus (BA 22)	L	−42	−23	+9	7603	−10.62
	R	+50	−20	+8	5647	−8.55

FIGURE 2 | Illustration of the contrast (thresholded at $q = 0.001$) between signed and spoken language comprehension overlaid on an individual human brain. Regions more active for spoken language comprehension (indicated in blue) include superior temporal gyrus in both hemispheres. Regions more active for signed language comprehension (in orange) include bilateral middle occipital cortex, bilateral pre-central gyrus, bilateral post central gyrus, bilateral IPS, left SPL, left SMG, and right pSTG. Up to 30 mm beneath the surface of cortex is displayed on the contrast map.

Not surprisingly, ASL comprehension engaged bilateral occipito-temporal cortex to a much greater extent than comprehension of audio-visual English. Activation in posterior middle temporal cortex (including area MT+) likely reflects perception of the much larger movements of the hands and arms produced within a larger physical space for sign language, compared to the

perception of the relatively small mouth movements of speech. Sign language movements also have a larger spatial frequency and thus are more likely to involve extra-foveal visual processes. Our findings replicate the results of other between-group studies that compared sign language comprehension by signers and audio-visual speech comprehension by hearing monolingual speakers, using relatively low-level baselines (e.g., MacSweeney et al., 2002a; Courtin et al., 2011).

Of particular interest is the extensive activation in bilateral parietal cortices observed for sign language comprehension relative to spoken language (see **Figure 2**). One partial explanation for greater parietal activation for ASL may lie in the semantic content of the sentences presented in the study—the sentences in both ASL and English conveyed information about the movement or static location of a referent. The ASL sentences involved classifier constructions in which locations in signing space correspond to referent locations and movements of the hand(s) through space depict the movements of a referent. Previous research has found that understanding this type of spatial language recruits left parietal cortex (the intraparietal sulcus) to a greater extent for ASL or BSL than for spoken English (MacSweeney et al., 2002b; McCullough et al., 2012). In addition, right parietal damage can impair comprehension of these types of spatial expressions (but does not cause sign aphasia), suggesting a critical role for the right hemisphere in comprehending spatial language in which physical space is used to express spatial concepts (e.g., Emmorey et al., 1995; Atkinson et al., 2005). In addition, the production of location and motion expressions using classifier constructions differentially recruits bilateral superior parietal cortex compared to the production of lexical signs (nouns) and compared to the production of lexical prepositions in spoken English (Emmorey et al., 2005, 2013).

Parietal cortex may also play a distinct role in phonological processing and working memory for sign language. Direct stimulation of the left supramarginal gyrus (SMG) results in handshape substitutions during picture naming (Corina et al., 1999), and MacSweeney et al. (2008b) reported greater activation in left SMG (extending into the superior parietal lobule) when deaf signers made phonological judgments about signs (do they share the same location?) than when they made phonological (rhyming) judgments about words, despite SMG engagement by both tasks relative to a baseline. Working memory for sign language also appears to engage parietal regions to a greater extent than for spoken language (Rönnberg et al., 2004; Buchsbaum et al., 2005; Bavelier et al., 2008; Pa et al., 2008). In particular, storage (the phonological buffer) and maintenance (rehearsal) of signs appear to rely more on parietal cortex compared to storage and maintenance processes for words.

Furthermore, the bilateral premotor and inferior parietal regions that were more active for sign than speech comprehension in Experiment 2 correspond to the predictive component of the Action Observation Network (AON), which is engaged when observing non-linguistic human body actions (e.g., Buccino et al., 2001; Caspers et al., 2010). The proposed function of the premotor-parietal (dorsal) component of the AON is the generation of predictions for observed manual actions (Kilner, 2011). Predictive coding accounts of the AON propose that premotor

and parietal cortices (the motor system used to produce manual actions) is active during action observation because it generates internal models that can be used to predict incoming visual input (Kilner, 2011; Schippers and Keysers, 2011). Premotor and parietal cortices are more engaged during active action understanding than during passive viewing of actions (Schippers and Keysers, 2011). Similarly, although several studies report prefrontal and parietal activation during active comprehension of signed sentences (e.g., Neville et al., 1998; MacSweeney et al., 2002a; Sakai et al., 2005) and single signs (e.g., MacSweeney et al., 2006), Emmorey et al. (2011) found little activation in these regions when deaf signers passively viewed strings of ASL signs. For sign language, this premotor-parietal circuit may be engaged in predicting the incoming visual input as part of active language comprehension.

Such a hypothesis is consistent with recent work by Pickering and Garrod (2013) who view language production and comprehension as forms of action and action perception, respectively. Applying forward modeling frameworks developed for human action to language comprehension, they propose that comprehenders use covert imitation and forward modeling to predict upcoming utterances. In this model, production and comprehension are integrated systems and both involve the extensive use of prediction. Perceivers of language construct forward models of others' linguistic actions that are based on their own potential actions. Thus, the differential premotor and parietal activation observed for sign language comprehension may be tied to the distinct neural substrate that supports sign language production.

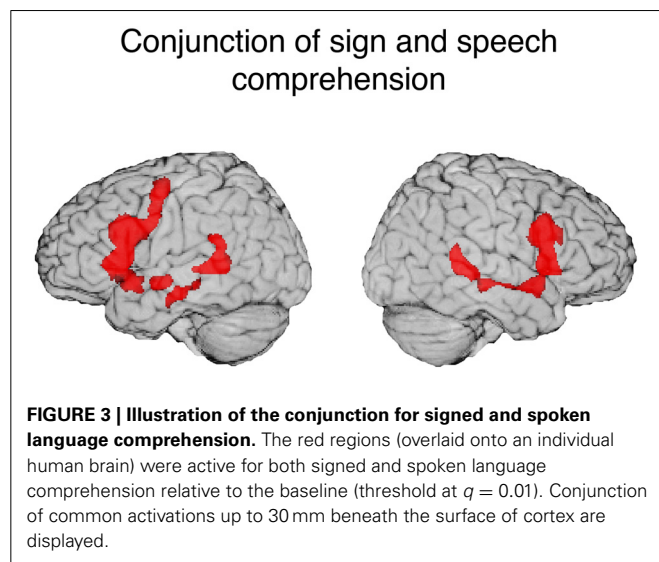## COMMON NEURAL SUBSTRATES FOR SIGNED vs. SPOKEN LANGUAGE COMPREHENSION

To identify shared neural substrates for sign and speech comprehension, we conducted a conjunction analysis for the contrast between each language and its baseline. The results are listed in **Table 3** and illustrated in **Figure 3**. Comprehension of both ASL and English engaged a bilateral fronto-temporal neural network, encompassing the inferior frontal gyrus (extending along the precentral gyrus in the left hemisphere) and the superior temporal sulcus (extending into posterior STG in the left hemisphere).

One striking result from the conjunction analysis is the degree to which modality independent activation during language comprehension is bilateral. Although reading print is highly left-lateralized, auditory and audio-visual spoken language comprehension engages a more bilateral network (e.g., Price, 2012). MacSweeney et al. (2002a) reported very similar bilateral fronto-temporal activation for comprehension of BSL sentences by native signers and comprehension of audiovisual English sentences by hearing native speakers. In contrast, Neville et al. (1998) observed left lateralized activation for hearing speakers reading English sentences, but bilateral activation for native signers comprehending ASL sentences. These findings highlight the importance of comparing sign language comprehension which always involves face-to-face interaction with the comprehension of audio-visual speech rather than with reading text or with a disembodied auditory-only speech signal (see also Hickok et al., 1998).

According to the dual stream model of speech processing proposed by Hickok and Poeppel (2007), phonological-level

Table 3 | Center of mass Talairach coordinates and cluster volumes for the conjunction of sign and spoken language comprehension (each vs. its baseline; thresholded at $q = 0.001$).

Region	Side	X	Y	Z	Vol mm³
FRONTAL CORTEX					
Inferior frontal gyrus (BA 45, 44, 4)	L	−38	+11	+19	13145
(BA 45, 44)	R	+42	+13	+17	7876
Medial frontal gyrus (BA 6)	L/R	+1	+3	+53	3286
TEMPORAL CORTEX					
Superior temporal sulcus (BA 22, 21)	L	−47	−8	−7	751
	L	−48	−41	+8	3486
	R	+50	−26	+1	4080
Hippocampus	L	−34	−19	−13	899
Parahippocampal gyrus (BA 36)	L	−19	−28	−14	99

Conjunction of sign and speech comprehension

FIGURE 3 | Illustration of the conjunction for signed and spoken language comprehension. The red regions (overlaid onto an individual human brain) were active for both signed and spoken language comprehension relative to the baseline (threshold at $q = 0.01$). Conjunction of common activations up to 30 mm beneath the surface of cortex are displayed.

processing and representation of speech is associated with middle to posterior portions of bilateral STS, with asymmetric functions in the left and right hemispheres. They suggest that left STS is more engaged in temporal and categorical processing of segment-level information, while right STS is more engaged in processing suprasegmental, prosodic information. Evidence that posterior STS might also be engaged in phonological processing for sign language comes from studies that examined linguistically structured pseudosigns. Pseudosigns have phonological structure for signers but do not access semantic or syntactic representations. A PET study by Petitto et al. (2000) found that viewing pseudosigns (as well as real signs) engaged STS bilaterally for deaf signers, but no activation was observed for hearing individuals who had not acquired a sign-based phonological system. Similarly, an fMRI study by Emmorey et al. (2011) reported that pseudosigns activated left posterior STS to a greater extent for deaf ASL signers than for hearing non-signers. Increased left posterior STS activation for signers is hypothesized to reflect heightened sensitivity to temporal body movements that conform to the phonological structure of ASL since dynamic movements

(e.g., path movements or changes in hand orientation) are critical to identifying syllabic structure in sign languages. Left STS may be significantly more active for signers than for non-signers because neurons in this region become particularly receptive to segment-level body movements that are linguistically structured and constrained. Right STS may also be engaged in sign-based phonological processing but perhaps only at the sentential level. An intriguing possibility is that left STS subserves categorical and combinatorial processing of sublexical sign structure, while right STS subserves more global phonological processes (e.g., sentential prosody expressed by movement; see Newman et al., 2010).

In addition, for spoken language, bilateral STS interfaces with MTG by mapping phonological representations onto lexical conceptual representations (the dorsal stream in the Hickok and Poeppel model). A similar interface may occur for signed language. The conjunction analysis revealed that STS activation extends into middle MTG for both ASL and English comprehension (see **Figure 3**). Results from a recent MEG study by Leonard et al. (2012) indicate that STS is engaged for both ASL signs and English words (in a sign/word picture matching task) during a relatively late time window associated with lexical-semantic processing (300–500 ms after stimulus onset), but only speech for hearing individuals activated STS during early sensory processing (80–120 ms). This finding suggests that STS activation for sign language is associated with lexical retrieval processes, rather than with early sensory processing which may be modality specific. Thus, bilateral activation in STS (extending into MTG) observed for both sign and speech may reflect amodal lexical-semantic and sublexical (phonological-level) processes (see Berent et al., 2013, for evidence for amodal phonological processes across signed and spoken languages).

Consistent with previous between-group studies, comprehension of audiovisual sentences and signed sentences both activate bilateral inferior frontal cortex, with activation extending anteriorly and dorsally in the left hemisphere. Comprehension functions associated with left inferior frontal cortex are numerous and are likely shared by both signed and spoken languages, e.g., syntactic processing, semantic retrieval, phonological-lexical integration—unification processes in Hagoort's (2013) model of language processing. Shared comprehension functions that may be associated with right inferior frontal cortex include prosodic processing and semantic inferencing (likely involved in the semantic anomaly detection task used here).

In sum, the conjunction results indicate that sign and audiovisual speech comprehension rely on a bilateral fronto-temporal network, with a slight left-hemisphere bias. The superior temporal sulcus is likely engaged in modality-independent phonological and lexical-semantic processes. Left inferior and middle frontal cortex may be engaged in various aspects of amodal syntactic, phonological, and semantic integration, while the right hemisphere homologue of Broca's area (BA 44/45) may be involved in semantic interpretation and sentence-level prosodic processing for both sign and speech comprehension.

SUMMARY AND CONCLUSIONS

The direct contrasts between ASL and English for both production and comprehension revealed relatively large differences in

neural resources related to perceptual and motor features of these two languages for hearing bimodal bilinguals (see **Figures 1, 2**). In contrast, direct contrasts between two spoken languages for unimodal bilinguals do not reveal such dramatic differences in neural activation (Gullberg and Indefrey, 2006). We suggest that the surface level differences between signed and spoken languages should not be dismissed as uninteresting and that these differences are critical for understanding how sensory-motor systems impact psycholinguistic processes and the underlying neural substrate for language.

A key psycholinguistic difference between signed and spoken language production is the role of perceptual (auditory or visual) feedback in monitoring language output and in learning new articulations for both adults and children (Emmorey et al., 2009a,c). Speakers use auditory feedback to detect errors (Postma and Noordanus, 1996) and to compare to "auditory targets" in the acquisition and maintenance of phonetic aspects of spoken segments and syllables (e.g., Guenther et al., 1998). Neural responses reflecting auditory feedback were observed in bilateral STC, which was significantly more engaged during speech than sign production as participants heard their own voices. In addition, we hypothesize that greater activation in right frontal cortex may reflect error detection processes involved in auditory monitoring of speech output as proposed within the DIVA modal of speech production (Tourville and Guenther, 2011).

For sign production, visual feedback cannot be easily parsed by the comprehension system (self-produced signs are not recognized very accurately; Emmorey et al., 2009a). In addition, "visual targets" are problematic for sign production because visual input from one's own signing differs substantially from visual input from another's signing. Thus, it is likely that signers rely on somatosensory more than visual feedback to monitor for errors and to acquire and maintain sign productions. Greater activation along the post-central gyri and anterior superior parietal cortex for signing compared to speaking may reflect somatosensory feedback received from signing.

The unexpected finding of increased activation in bilateral occipital cortex for speaking compared to signing may also be related to differences in sensory feedback. Specifically, we hypothesize that greater occipital activation for speaking may actually be due to suppression of cortical activity during signing. We speculate that cortical attenuation in visual cortex may serve to distinguish between visual stimulation arising from the signer's own movements and externally produced movements toward the body and face. Predicting the visual consequences of one's own actions may attenuate activation in visual cortex, which would help to dissociate sensory signals generated by one's own actions from sensory signals that are externally generated by the environment.

An important finding from these studies is that *both* sign language production and comprehension engaged parietal cortex to a greater extent than spoken language. In fact, the peak coordinates within the anterior superior parietal lobule (in the post-central sulcus) are within 10 mm of each other for sign production (-35, -31, $+50$; $+33$, -35, $+51$) and sign comprehension (-27, -46, $+44$; $+30$, -46, $+44$). We hypothesize that anterior SPL (possibly in conjunction with inferior parietal

cortex) is more engaged during sign language comprehension because the *production* system for signing differs from speech and that production and comprehension are interweaved for sign language, as has been proposed for spoken language (Pickering and Garrod, 2013). Specifically, parietal regions may be involved in creating a forward model that predicts the incoming visual manual signal during comprehension. Recently, Hosemann et al. (2013) provided ERP evidence suggesting that sentence comprehension in sign language (in this case, German Sign Language) involves the use of forward modeling such that manual information in a transitional movement is used to predict an upcoming sign. Consistent with these results, the MEG study by Leonard et al. (2012) reported a larger response in left parietal cortex (in the intraparietal sulcus) for incongruent signs (those that did not match a preceding picture) than congruent signs, but no difference in parietal cortex was observed for spoken language. Thus, evidence is mounting that parietal cortex may be involved in internal simulations (generating predictions) during sign language comprehension. Internal simulations differ between spoken and sign languages because their production systems involve different articulators.

The conjunction results point to neural substrates that support modality-independent, shared computational processes for spoken and signed languages. The conjunction studies for production that were reviewed here along with the results from the comprehension conjunction (**Figure 3**) identify left inferior frontal cortex as a key amodal language area. For production, the left pre-central gyrus and the supplementary motor area have been found to be jointly engaged when signing or speaking (and when covertly signing or speaking in a rehearsal task—see Pa et al., 2008), pointing to an amodal role in the complex articulations required by the human language system. For comprehension, bilateral STS was engaged for both English and ASL, and we hypothesize there may be similar asymmetric functions for left and right STS for both language types. Left anterior STS regions may be engaged in amodal syntactic processes (e.g., Friederici et al., 2003), while posterior STS regions may be involved in lexical-phonological processes that are independent of modality. Right STS may function to integrate suprasegmental, prosodic information conveyed either by vocal intonation or intonation expressed by facial expressions and manual prosody (see Sandler, 1999; Dachovsky and Sandler, 2009, for evidence of visual prosody in sign languages). Right inferior frontal cortex was also engaged during spoken and sign language comprehension and may be involved in semantic processing, as well as prosodic segmentation during sentence comprehension (cf. Friederici, 2011).

In sum, results from direct contrasts between signing and speaking and between visual and audio-visual language comprehension revealed non-obvious distinctions between the two language types. The differences between sign and speech were not restricted to input/output differences in primary sensory and motor systems—surface level differences were also observed in heteromodal association cortex, suggesting that higher order systems may be needed to integrate modality-specific information. Our conjunction analysis revealed the expected overlap in left perisylvian language regions but also indicated an important role for the right hemisphere during face-to-face

language comprehension. Further detailed studies that target specific linguistic processes are needed to identify invariant structure-function associations within the language network and to demarcate the specific functional roles of cortical regions that distinguish between languages by hand and languages by mouth.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://www.frontiersin.org/journal/10.3389/fpsyg.2014.00484/abstract

REFERENCES

Atkinson, J., Marshall, J., Woll, B., and Thacker, A. (2005). Testing comprehension abilities in users of British Sign Language following CVA. *Brain Lang.* 94, 233–248. doi: 10.1016/j.bandl.2004.12.008

Bates, E., D'Amico, S., Jacobsen, T., Szekely, A., Andonova, E., Devescovi, A., et al. (2003). Timed picture naming in seven languages. *Psychon. Bull. Rev.* 10, 344–380. doi: 10.3758/BF03196494

Bavelier, D., Brozinsky, C., Tomann, A., Mitchell, T., Neville, H., and Lui, G. (2001). Impact of deafness and early exposure to sign language on cerebral organization for motion processing. *J. Neurosci.* 21, 8931–8942.

Bavelier, D., Newman, A. J., Mukherjerr, M., Hauser, P., Kemeny, S., Braun, A., et al. (2008). Encoding, rehearsal, and recall in signers and speakers: shared network but differential engagement. *Cereb. Cortex* 18, 2263–2274. doi: 10.1093/cercor/bhm248

Berent, I., Dupuis, A., and Brentari, D. (2013). Amodal aspects of linguistic design. *PloS ONE* 8:e60617. doi: 10.1371/journal.pone.0060617

Bosworth, R., and Dobkins, K. (2001). Visual field asymmetries for motion processing in deaf and hearing signers. *Brain Cogn.* 49, 170–181. doi: 10.1006/brcg.2001.1498

Braun, A. R., Guillemin, A., Hosey, L., and Varga, M. (2001). The neural organization of discourse: an $H_2^{15}O$-PET study of narrative production in English and American Sign Language. *Brain* 124, 2028–2044. doi: 10.1093/brain/124.10.2028

Buccino, G., Binkofski, F., Fink, G. R., Fadiga, L., Fogassi, L., Gallese, V., et al. (2001). Action observation activates premotor and parietal areas in a somatotopic manner: an fMRI study. *Eur. J. Neurosci.* 13, 400–404. doi: 10.1111/j.1460-9568.2001.01385.x

Buchsbaum, B., Pickell, B., Love, T., Hatrak, M., Bellugi, U., and Hickok, G. (2005). Neural substrates for working memory in deaf signers: fMRI study and lesion case report. *Brain Lang.* 95, 265–272. doi: 10.1016/j.bandl.2005.01.009

Cardin, V., Orfanidou, E., Rönnberg, J., Capek, C. M., Rudner, M., and Woll, B. (2013). Dissociating cognitive and sensory neural plasticity in human superior temporal cortex. *Nat. Commun.* 4, 1473. doi: 10.1038/ncomms2463

Caspers, S., Zilles, K., Laird, A. R., and Eickhoff, S. B. (2010). ALE meta-analysis of action observation and imitation in the human brain. *Neuroimage* 50, 1148–1167. doi: 10.1016/j.neuroimage.2009.12.112

Chen, G., Saad, Z., Nath, A. R., Beauchamp, M. S., and Cox, R. W. (2012). FMRI group analysis combining effect estimates and their variances. *Neuroimage* 60, 747–765. doi: 10.1016/j.neuroimage.2011.12.060

Corina, D. P., and Knapp, H. (2006). Sign language processing and the mirror neuron system. *Cortex* 42, 529–539. doi: 10.1016/S0010-9452(08)70393-9

Corina, D. P., Lawyer, L. A., and Cates, D. (2013). Cross-linguistic differences in the neural representation of human language: evidence from users of signed languages. *Front. Psychol.* 3:587. doi: 10.3389/fpsyg.2012.00587

Corina, D. P., McBurney, S. L., Dodrill, C., Hinshaw, K., Brinkley, J., and Ojemann, G. (1999). Functional roles of Broca's area and SMG: evidence from cortical stimulation mapping in a deaf signer. *Neuroimage* 10, 570–581. doi: 10.1006/nimg.1999.0499

Corina, D. P., San Jose-Robertson, L., Guillemin, A., High, J., and Braun, A. R. (2003). Language lateralization in a bimanual language. *J. Cogn. Neurosci.* 15, 718–730. doi: 10.1162/jocn.2003.15.5.718

Courtin, C., Jobard, G., Vigneau, M., Beaucousin, V., Razafimandimby, A., Hervé, P. Y., et al. (2011). A common neural system is activated in hearing non-signers to process French Sign language and spoken French. *Brain Res. Bull.* 84, 75–87. doi: 10.1016/j.brainresbull.2010.09.013

Cox, R. W. (1996). AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Comput. Biomed. Res.* 29, 162–173. doi: 10.1006/cbmr.1996.0014

Creem-Regehr, S. (2009). Sensory-motor and cognitive function of the human posterior parietal cortex involved in manual actions. *Neurobiol. Learn. Mem.* 91, 166–171. doi: 10.1016/j.nlm.2008.10.004

Dachovsky, S., and Sandler, W. (2009). Visual intonation in the prosody of a sign language. *Lang. Speech* 52, 287–314. doi: 10.1177/0023830909103175

Damasio, H., Grabowski, T. J., Frank, R., Knosp, B., Hichwa, R. D., Watkins, G. L., et al. (1994). "PET-Brainvox, a technique for neuroanatomical analysis of positron emission tomography images," in *Quantification of Brain Function*, eds K. Uemura, N. A. Lassen, T. Jones, and I. Kanno (Amsterdam: Elsevier), 465–474.

Emmorey, K. (2002). *Language, Cognition, and the Brain: INSIGHTS from Sign Language Research*. Mahwah, NJ: Lawrence Erlbaum and Associates.

Emmorey, K., Bosworth, R., and Kraljic, T. (2009a). Visual feedback and self-monitoring of sign language. *J. Mem. Lang.* 61, 398–411. doi: 10.1016/j.jml.2009.06.001

Emmorey, K., Corina, D., and Bellugi, U. (1995). "Differential processing of topographic and referential functions of space" in *Language, Gesture, and Space*, eds K. Emmorey and J. Reilly, (Hillsdale, NJ: Lawrence Erlbaum Associates), 43–62.

Emmorey, K., Gertsberg, N., Korpics, F., and Wright, C. E. (2009b). The influence of visual feedback and register changes on sign language production: a kinematic study with deaf signers. *Appl. Psycholinguist.* 30, 187–203. doi: 10.1017/S0142716408090085

Emmorey, K., Grabowski, T., McCullough, S., Damasio, H., Ponto, L., Hichwa, R., et al. (2003). Neural systems underlying lexical retrieval for sign language. *Neuropsychologia* 41, 85–95. doi: 10.1016/S0028-3932(02)00089-1

Emmorey, K., Grabowski, T. J., McCullough, S., Ponto, L., Hichwa, R., and Damasio, H. (2005). The neural correlates of spatial language in English and American Sign Language: a PET study with hearing bilinguals. *Neuroimage* 24, 832–840. doi: 10.1016/j.neuroimage.2004.10.008

Emmorey, K., Korpics, F., and Petronio, K. (2009c). The use of visual feedback during signing: evidence from signers with impaired vision. *J. Deaf Stud. Deaf Educ.* 14, 99–104. doi: 10.1093/deafed/enn020

Emmorey, K., and McCullough, S. (2009). The bimodal brain: effects of sign language experience. *Brain Lang.* 110, 208–221. doi: 10.1016/j.bandl.2008.03.005

Emmorey, K., McCullough, S., Mehta, S., Ponto, L. L. B., and Grabowski, T. J. (2013). The biology of linguistic expression impacts neural correlates for spatial language. *J. Cogn. Neurosci.* 25, 517–533. doi: 10.1162/jocn_a_00339

Emmorey, K., Mehta, S., and Grabowski, T. J. (2007). The neural correlates of sign versus word production. *Neuroimage* 36, 202–208. doi: 10.1016/j.neuroimage.2007.02.040

Emmorey, K., Xu, J., and Braun, A. (2011). Neural responses to meaningless pseudosigns: evidence for sign-based phonetic processing in superior temporal cortex. *Brain Lang.* 117, 34–38. doi: 10.1016/j.bandl.2010.10.003

Finney, E. M., Fine, I., and Dobkins, K. R. (2001). Visual stimuli activate auditory cortex in the deaf. *Nat. Neurosci.* 4, 1171–1173. doi: 10.1038/nn763

Frank, R. J., Damasio, H., and Grabowski, T. J. (1997). Brainvox: an interactive, multimodal visualization and analysis system for neuroanatomical imaging. *Neuroimage* 5, 13–30. doi: 10.1006/nimg.1996.0250

Friederici, A. D. (2011). The brain basis of language processing: from structure to function. *Physiol. Rev.* 91, 1357–1392. doi: 10.1152/physrev.00006.2011

Friederici, A. D., Rüschemyer, S. A., Hahne, A., and Fiebach, C. J. (2003). The role of left inferior frontal and superior temporal cortex in sentence comprehension: localizing syntactic and semantic processes. *Cereb. Cortex* 13, 170–177. doi: 10.1093/cercor/13.2.170

Friston, K. J., Holmes, A. P., Worsley, K. J., Poline, J.-B., Frith, C. D., and Frackowiak, R. S. J. (1995). Statistical parametric maps in functional imaging: a general linear approach. *Hum. Brain Mapp.* 2, 189–210. doi: 10.1002/hbm.460020402

Grabowski, T. J., Damasio, H., Frank, R., Hichwa, R. D., Ponto, L. L., and Watkins, G. L. (1995). A new technique for PET slice orientation and MRI-PET coregistration. *Hum. Brain Mapp.* 2, 123–133. doi: 10.1002/hbm.460020302

Grabowski, T. J., Frank, R. J., Brown, C. K., Damasio, H., Ponto, L. L. B., Watkins, G. L., et al. (1996). Reliability of PET activation across statistical methods, subject groups, and sample sizes. *Hum. Brain Mapp.* 4, 227–239. doi: 10.1002/(SICI)1097-0193(1996)4:1<23::AID-HBM2>3.0.CO;2-R

Grabowski, T. J., Frank, R. J., Szumski, N. R., Brown, C. K., and Damasio, H. (2000). Validation of partial tissue segmentation of single-channel magnetic resonance images. *Neuroimage* 12, 640–656. doi: 10.1006/nimg.2000.0649

Grodzinsky, Y., and Amunts, K. (2005). *Broca's Region.* Oxford: Oxford University Press.

Guenther, F. H., Hampson, M., and Johnson, D. (1998). A theoretical investigation of reference frames for the planning of speech movements. *Psychol. Rev.* 105, 611–633. doi: 10.1037/0033-295X.105.4.611-633

Gullberg, M., and Indefrey, P. (2006). *The Cognitive Neuroscience of Second Language Acquisition.* Oxford, UK: Blackwell Publishing.

Hagoort, P. (2013). MUC (Memory, Unification, Control) and beyond. *Front. Psychol.* 4:416. doi: 10.3389/fpsyg.2013.00416

Herscovitch, P., Markham, J., and Raichle, M. E. (1983). Brain blood flow measure with intravenous H_2^{15} O. I. Theory and error analysis. *J. Nucl. Med.* 24, 782–789.

Hesse, M. D., Nishitani, N., Fink, G. R., Jousmäki, V., and Hari, R. (2010). Attenuation of somatosensory activation for self-produced tactile stimulation. *Cereb. Cortex* 20, 425–432. doi: 10.1093/cercor/bhp110

Hichwa, R. D., Ponto, L. L., and Watkins, G. L. (1995). "Clinical blood flow measurement with [150]water and positron emission tomography (PET)" in *Symposium Proceedings of the International Symposium on Chemists' Views of Imaging Centers,* ed A. M. Emran (New York, NY: Plenum Publishing).

Hickok, G., Bellugi, U., and Klima, E. (1998). What's right about the neural organization of sign language? A perspective on recent neuroimaging studies. *Trends Cogn. Sci.* 2, 465–468. doi: 10.1016/S1364-6613(98)01263-7

Hickok, G., Bellugi, U., and Klima, E. S. (1996). The neurobiology of sign language and its implications for the neural basis of language. *Nature* 381, 699–702. doi: 10.1038/381699a0

Hickok, G., Okada, K., and Serences, J. T. (2009). Area Spt in the human planum temporale supports sensory-motor integration for speech processing. *J. Neurophysiol.* 101, 2725–2732. doi: 10.1152/jn.91099.2008

Hickok, G., and Poeppel, D. (2007). The cortical organization of speech processing. *Nat. Rev. Neurosci.* 8, 393–402. doi: 10.1038/nrn2113

Horwitz, B., Amunts, K., Bhattacharyya, R., Patkin, D., Jeffries, K., Zilles, K., et al. (2003). Activation of Broca's area during the production of spoken and signed langauge: a combined cytoarchitectonic mapping and PET analysis. *Neuropsychologia* 41, 1868–1876. doi: 10.1016/S0028-3932(03)00125-8

Hosemann, J., Hermann, A., Steinbach, M., Bornkessel-Schlesewsky, I., and Schlesewsky, M. (2013). Lexical prediction via forward models: N400 evidence from German Sign Language. *Neuropsychologia* 51, 2224–2237. doi: 10.1016/j.neuropsychologia.2013.07.013

Houde, J. F., Nagarajan, S. S., Sekihara, K., and Merzenich, M. M. (2002). Modulation of auditory cortex during speech: an MEG study. *J. Cogn. Neurosci.* 14, 112–1138. doi: 10.1162/089892902760807140

Hu, Z., Wang, W., Liu, H., Peng, D., Yang, Y., Li, K., et al. (2011). Brain activations associated with sign production using word and picture inputs in deaf signers. *Brain Lang.* 116, 64–70. doi: 10.1016/j.bandl.2010.11.006

Indefrey, P., and Levelt, W. J. M. (2004). The spatial and temporal signatures of word production components. *Cognition* 92, 101–144. doi: 10.1016/j.cognition.2002.06.001

Jenkinson, M., Bannister, P. R., Brady, J. M., and Smith, S. M. (2002). Improved optimisation for the robust and accurate linear registration and motion correction of brain images. *Neuroimage* 17, 825–841. doi: 10.1006/nimg.2002.1132

Jenkinson, M., and Smith, S. M. (2001). A global optimisation method for robust affine registration of brain images. *Med. Image Anal.* 5, 143–156. doi: 10.1016/S1361-8415(01)00036-6

Kassubek, J., Hickok, G., and Erhard, P. (2004). Involvement of classical anterior and posterior language areas in sign language production, as investigated by 4 T functional magnetic resonance imaging. *Neurosci. Lett.* 364, 168–172. doi: 10.1016/j.neulet.2004.04.088

Kilner, J. M. (2011). More than one pathway to action understanding. *Trends Cogn. Sci.* 15, 352–357. doi: 10.1016/j.tics.2011.06.005

Leonard, M. K., Ramirez, N. F., Torres, C., Travis, K. E., Hatrak, M., Mayberry, R. I., et al. (2012). Signed words in the congenitally deaf evoke typical late lexicosemantic responses with no early visual responses in left superior temporal cortex. *J. Neurosci.* 32, 9700–9705. doi: 10.1523/JNEUROSCI.1002-12.2012

Lorey, B., Pilgramm, S., Walter, B., Stark, R., Munzert, J., and Zentgraf, K. (2010). Your mind's hand: motor imagery of pointing movements with different accuracy. *Neuroimage* 49, 3239–3247. doi: 10.1016/j.neuroimage.2009.11.038

MacSweeney, M., Campbell, R., Woll, B., Brammer, M. J., Giampietro, V., David, A. S., et al. (2006). Lexical and sentential processing in British Sign Language. *Hum. Brain Mapp.* 27, 63–76. doi: 10.1002/hbm.20167

MacSweeney, M., Capek, C. M., Campbell, R., and Woll, B. (2008a). The signing brain: the neurobiology of sign language. *Trends Cogn. Sci.* 12, 432–440. doi: 10.1016/j.tics.2008.07.010

MacSweeney, M., Waters, D., Brammer, M. J., Woll, B., and Goswami, U. (2008b). Phonological processing in deaf signers and the impact of age of first acquisition. *Neuroimage* 40, 1369–1379. doi: 10.1016/j.neuroimage.2007.12.047

MacSweeney, M., Woll, B., Campbell, R., Calvert, G. A., McGuire, P. K., David, A. S., et al. (2002b). Neural correlates of British Sign Language comprehension: spatial processing demands of topographic language. *J. Cogn. Neurosci.* 14, 1604–1075. doi: 10.1162/089892902320474517

MacSweeney, M., Woll, B., Campbell, R., McGuire, P. K., David, A. S., Williams, S. C. et al. (2002a). Neural systems underlying British Sign Language and audio-visual English processing in native users. *Brain* 125, 1583–1593. doi: 10.1093/brain/awf153

McCullough, S., Saygin, A. P., Korpics, F., and Emmorey, K. (2012). Motion-sensitive cortex and motion semantics in American Sign Language. *Neuroimage* 63, 111–118. doi: 10.1016/j.neuroimage.2012.06.029

Miller, A. J. (1991). Least squares routines to supplement those of gentleman. *Appl. Statist. Algorithm* 274, 458–478.

Neville, H. J., Bavelier, D., Corina, D., Rauschecker, J., Karni, A., Lalwani, A., et al. (1998). Cerebral organization for language in deaf and hearing subjects: biological constraints and effects of experience. *Proc. Natl. Acad. Sci. U.S.A.* 95, 992–929. doi: 10.1073/pnas.95.3.922

Newman, A. J., Supalla, T., Hauser, P. C., Newport, E. L., and Bavelier, D. (2010). Prosodic and narrative processing in American Sign Language: an fMRI study. *Neuroimage* 52, 669–676. doi: 10.1016/j.neuroimage.2010.03.055

Nichols, T., Brett, M., Andersson, J., Wager, T., and Poline, J. B. (2005). Valid conjunction inference with the minimum statistic. *Neuroimage* 25, 653–660. doi: 10.1016/j.neuroimage.2004.12.005

Numminen, J., Salmelin, R., and Hari, R. (1999). Subject's own speech reduces reactivity of the human auditory cortex. *Neurosci. Lett.* 265, 119–122. doi: 10.1016/S0304-3940(99)00218-9

Pa, J., Wilson, S. M., Pickell, H., Bellugi, U., and Hickok, G. (2008). Neural organization of linguistic short-term memory is sensory modality-dependent: evidence from signed and spoken language. *J. Cogn. Neurosci.* 20, 2198–2210. doi: 10.1162/jocn.2008.20154

Parker Jones, 'Ö., Green, D. W., Grogan, A., Pliatskias, C., Filippopolitis, K., Ali, N., et al. (2012). Where, when and why brain activation differs for bilinguals and monolinguals during picture naming and reading aloud. *Cereb. Cortex* 22, 892–902. doi: 10.1093/cercor/bhr161

Petitto, L. A., Zatorre, R. J., Gauna, K., Nikelski, E. J., Dostie, D., and Evans, A. C. (2000). Speech-like cerebral activity in profoundly deaf people processing signed languages: implications for the neural basis of human language. *Proc. Natl. Acad. Sci. U.S.A.* 97, 13961–13966. doi: 10.1073/pnas.97.25.13961

Pickering, M. J., and Garrod, S. (2013). An integrated theory of language production and comprehension. *J. Behav. Brain Sci.* 36, 329–392. doi: 10.1017/S0140525X12001495

Postma, A., and Noordanus, C. (1996). Production and detection of speech errors in silent, mouthed, noise-masked, and normal auditory feedback speech. *Lang. Speech* 39, 375–392.

Price, C. J. (2012). A review and synthesis of the first 20 years of PET and fMRI studies of heard speech, spoken language and reading. *Neuroimage* 62, 816–847. doi: 10.1016/j.neuroimage.2012.04.062

Rönnberg, J., Rudner, M., and Ingvar, M. (2004). Neural correlates of working memory for sign language. *Cogn. Brain Res.* 20, 165–182. doi: 10.1016/j.cogbrainres.2004.03.002

Sakai, K. L., Tatsuno, Y., Suzuki, K., Kimura, H., and Ichida, Y. (2005). Sign and speech: amodal commonality in left hemisphere dominance for comprehension of sentences. *Brain* 128, 1407–1417. doi: 10.1093/brain/awh465

Sandler, W. (1999). Prosody in two natural language modalities. *Lang. Speech* 42, 127–142. doi: 10.1177/00238309990420020101

Saygin, A. P., McCullough, S., Alac, M., and Emmorey, K. (2010). Modulation of BOLD response in motion-sensitive lateral temporal cortex by real and fictive motion sentences. *J. Cogn. Neurosci.* 22, 2480–2490. doi: 10.1162/jocn.2009.21388

Schippers, M. B., and Keysers, C. (2011). Mapping the flow of information within the putative mirror neuron sysem during gesture observation. *Neuroimage* 57, 37–44. doi: 10.1016/j.neuroimage.2011.02.018

Simmons, A. J., Wise, R. J. S., and Leech, R. (2011). Two tongues, one brain: imaging bilingual speech production. *Front. Psychol.* 2:166. doi: 10.3389/fpsyg.2011.00166

Söderfelt, B., Ingvar, M., Rönnberg, J., Eriksson, L., Serrander, M., and Stone-Elander, S. (1997). Signed and spoken language perception studied by positron emission tomography. *Neurology* 49, 82–87. doi: 10.1212/WNL.49.1.82

Tourville, J. A., and Guenther, F. H. (2011). The DIVA model: a neural theory of speech acquisition and production. *Lang. Cogn. Process.* 26, 952–981. doi: 10.1080/01690960903498424

Tourville, J. A., Reilly, K. J., and Guenther, F. H. (2008). Neural mechanisms controlling auditory feedback control of speech. *Neuroimage* 39, 1429–1443. doi: 10.1016/j.neuroimage.2007.09.054

Tungaraza, R. F., Mehta, S., Emmorey, K., and Grabowski, T. J. (2011). "White matter differences between deaf ASL signers and hearing English speakers," in *Poster Presented at Human Brain Mapping* (Québec city, QC).

Woods, R. P., Dapretto, M., Sicotte, N. L., Toga, A. W., and Mazziotta, J. C. (1999). Creation and use of a Talairach-compatible atlas for accurate, automated, non-linear intersubject registration, and analysis of functional imaging data. *Hum. Brain Mapp.* 8, 73–79. doi: 10.1002/(SICI)1097-0193(1999)8:2/33.0.CO;2-7

Worsley, K. J. (1994). Local maxima and the expected Euler characteristic of excursion sets of chi-squared, F and t fields. *Adv. Appl. Probabil.* 26, 13–42. doi: 10.2307/1427576

Worsley, K. J., Evans, A. C., Marrett, S., and Neelin, P., (1992). A three dimensional statistical analysis for CBF activation studies in human brain. *J. Cereb. Blood Flow Metab.* 12, 900–918. doi: 10.1038/jcbfm.1992.127

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

On language acquisition in speech and sign: development of combinatorial structure in both modalities

*Gary Morgan**

Language and Communication Science, City University London, London, UK

Languages are composed of a conventionalized system of parts which allow speakers and signers to generate an infinite number of form-meaning mappings through phonological and morphological combinations. This level of linguistic organization distinguishes language from other communicative acts such as gestures. In contrast to signs, gestures are made up of meaning units that are mostly holistic. Children exposed to signed and spoken languages from early in life develop grammatical structure following similar rates and patterns. This is interesting, because signed languages are perceived and articulated in very different ways to their spoken counterparts with many signs displaying surface resemblances to gestures. The acquisition of forms and meanings in child signers and talkers might thus have been a different process. Yet in one sense both groups are faced with a similar problem: "how do I make a language with combinatorial structure"? In this paper I argue first language development itself enables this to happen and by broadly similar mechanisms across modalities. Combinatorial structure is the outcome of phonological simplifications and productivity in using verb morphology by children in sign and speech.

Keywords: sign, acquisition, phonology, classifiers, componential structure

INTRODUCTION

Signed languages have all the levels of complexity and expressive power as spoken languages, they are processed in similar ways by cognitive and related brain networks (Emmorey, 2002) and they can be acquired as native languages by children following the same developmental stages as those identified for spoken language (Petitto, 1997; Chamberlain et al., 2000; Morgan and Woll, 2002; Baker and Woll, 2009; Chen Pichler, 2012). Native signers are a rare group, as only 5–10% of deaf children have deaf parents (Mitchell and Karchmer, 2004). In this paper I focus on two main issues in native sign language acquisition: (1) the relationship between gestures and signs and (2) the emergence of combinatorial structure during language development. To illustrate both issues I use case studies of native signers acquiring BSL. I ague that combinatorial structure distinguishes signs and gestures, and that this difference comes about because of language acquisition mechanisms.

The paper is organized as follows: the first section focuses on signs and gestures and explores how these two forms of semiotic communication are different, based on the presence or absence of conventionalized linguistic representations. In section 2, I describe how children's development of language leads to combinatorial structure. In section 3, I illustrate the points made in the first two sections by reviewing firstly, the linguistic organization and acquisition of phonology in native signers. The intention is to bring out the broad similarities that exist for phonological development across modalities. Then secondly, in section 4, I describe how spatial utterances in signed and spoken language are organized linguistically and develop in native signers of BSL. This development illustrates the difference between holistic gestures and conventionalized signing with respect to combinatorial structure and also why

productivity is important. The paper concludes with some discussion of how research on child sign learners can contribute to a greater understanding of language acquisition in general.

THE DIFFERENCE BETWEEN SIGNS AND GESTURES IS BASED ON LINGUISTIC ORGANIZATION

The inclusion of the languages of deaf communities into linguistic, psychological, and neurological research has enriched these disciplines and revealed which cognitive processes deal primarily with speech versus those devoted to cross-modality instances of language (Klima and Bellugi, 1979; Pfau et al., 2012). We see that sign languages have a linguistic organization of form and meaning components following traditional ideas of recursion and hierarchical patterning present in all human languages (Chomsky, 1965). These qualities do not appear in gestures which when produced with speech are mostly dependent on the spoken language system (e.g., de Ruiter, 2000). Although quite rare in everyday communication, when gestures are articulated in the absence of speech they take on more language-like properties (Goldin-Meadow et al., 1996). Kendon's (2004) continuum positions gesture and sign language a distance from each other (McNeill, 1992) but a continuum indicates quantitative rather than qualitative differences (i.e., gesture and sign have a similar and contiguous semiotic underpinning) but this is a contentious point (Singleton et al., 1995). The dis/continuity debate appears in several of the following sections.

It is probably the case that the start of how signs began to evolve comes from homesign systems used before deaf people came together in schools (Senghas et al., 2004; Brentari et al., 2012). In this account, the evolutionary beginnings of specific classes of signs (e.g., classifiers) may have developed from gestures

of the surrounding spoken language community (Duncan, 2003; Zeshan, 2003; Van Loon et al., in press).

Although gestures share many similarities with sign languages, on linguistic grounds gestures remain holistic, gradient, and not decomposable (McNeill, 1992; McNeill and Duncan, 2000; Kendon, 2004). In terms of linguistic representation gestures lack the combinatorial structure present in sign language phonology and morphology.

These differences between gestures and signs are important for the field of sign language acquisition where the role of gesture in sign development is still not clearly understood (Volterra and Erting, 1994; Schick, 2004). While the transition between gestures and words in spoken language development has been well documented (Goldin-Meadow, 2003), it has been more difficult to track in signing children because they happen in the same modality. When young hearing children use gestures as they acquire their native spoken language it is transparent to see how the two systems are being used separately or together (Volterra and Erting, 1994). It is less clear how this happens during signed language acquisition or even in the adult system itself (Liddell, 2003). For language development research the debate is about the dis/continuity between gestures and signs. This can be subsumed into a larger question about the modularity or cognitive generality of language (e.g., Petitto, 1987). Returning to the continuum between gestures and sign languages, our question is do children adapt their gestures into sign languages as a gradual and continual process or does sign language acquisition lead to a qualitative reorganization of gesture? The latter implies a profound impact of child language acquisition mechanisms on the structure of language rather than language emerging from gesture as a diachronic process over time. For a similar debate see the role of children in the field of creole genesis (de Graff, 1999). In the following sections I argue that during language development native signers turn communicative gestures into combinatorial grammar.

CHILDREN'S ACQUISITION OF LANGUAGE LEADS TO COMBINATORIAL STRUCTURE

It has long been suggested that *development* itself drives the change between holistic gestures and combinatorial signs (Newport, 1990). This can be observed in different scenarios. For example by the individual homesigner who creates a conventional system over a life-time (i.e., morphology in Goldin-Meadow et al., 1995), in studies of signed language evolution, where each successive cohort shapes the language from isolated homesigns as a substrate onto a conventional grammar (Senghas et al., 2004; Sandler et al., 2005; Brentari et al., 2012). A still further example is the focus of the current paper which is how the native acquisition of a signed language brings about conventional patterns of phonological and morphological structures. When signing children start to communicate they use communicative gestures, as hearing children do, but at some point in their acquisition of a language they arrive at a system of phonological and morphological conventions. It might be that gestures and sign differ in their linguistic status because native signers are able to create combinatorial language (Singleton and Newport, 2004).

But in one sense conventionalization occurs in every single child who learns a language from their care-givers (e.g., Valian, 2009). Most hearing and deaf native signing children experience optimal input but they still need to arrive at a conventionalized linguistic representation which approximates the adult model. In the bulk of this paper I look at how communication becomes sign language and takes on the linguistic properties underlying a phonological and morphological system. Previous research has documented in detail how gesture and sign differ, however, in this paper I attempt to provide a unifying rationale for how holistic gestures become combinatorial grammar through language development. The argument is that combinatorial structure is an outcome of some well-known mechanisms inherent to first language development: a set of phonological processes in word/sign production (Smith, 1973) and achieving productivity in a morphological system for sign/spoken constructions (Brown, 1973; Tomasello, 1992). I argue that this is why sign languages have the structures they have and come from but are distinguished from gesture.

THE LINGUISTIC ORGANIZATION OF PHONOLOGY AND SIMPLIFICATION PROCESSES ACROSS SIGN AND SPEECH DEVELOPMENT

The first linguistic descriptions of American Sign Language (ASL) by Stokoe (1960) and Klima and Bellugi (1979) demonstrated a duality of patterning (i.e., control of a phonology and grammar). Individual signs could be broken down into handshape, movement and location parameters, demonstrating systematic phonological structure: a hallmark of all human language (and a contrast to gestures). Later research extended this to other natural signed languages, e.g., British (BSL) or Catalan Sign Language (LSC) and eventually signing was described using mainstream phonological and linguistic theories (e.g., Sandler and Lillo-Martin, 2006).

One part of phonological structure is the existence of minimal pairs, where two lexical items differ by one phoneme only, e.g., [ki] vs. [ti] in "key and tea." Two signs can also differ in only one parameter, e.g., the BSL signs NAME and AFTERNOON have the same handshape and movement, but the hand moves from the forehead location in NAME and the chin in AFTERNOON. Not all exemplars in each of the sign parameters have the same level of complexity in their internal representations. For example, in the movement parameter a simple exemplar would be in a sign with a straight trajectory of the hand while a sign with a complex movement would be one where the hands move both internally (e.g., by opening and closing) as the hand moves in an arc shape. Brentari (1998) argues these different complexities are related to markedness. This concept has many different interpretations within linguistic theories and is the subject of much debate. While many linguists explicitly define markedness as a grammatical force (i.e., constraints), others have equated it with a notion of sensorimotor complexity (not necessarily specific to language or the grammar). This last interpretation is used in this paper when discussing language acquisition processes.

Brentari (1998) makes clear predictions for which handshapes are marked in ASL, based on the number of features present in the phonological representation, e.g., simultaneous extension and

flexion of fingers is one more feature in the representation than one of these movements in isolation. The most complex signs also have the least commonly occurring parameter types in the language. For example, a handshape containing the largest number of selected features in its representation appears in the fewest number of signs. Conversely "unmarked" phonological parameters are those that are phonetically and phonologically simple, as well as frequently occurring in the language. Markedness is important in studies of native signer's first signs.

THE EMERGENCE OF PHONOLOGY IS SIMILAR ACROSS MODALITIES

Turning first to the vocal modality, hearing children establish phonological representations for the input extremely early in development (Jusczyk et al., 2002) but in production it is difficult to distinguish actual first words from canonical babbling sounds. First words are often described as unanalyzed or frozen forms (akin to gestures) rather than generated from a system of individual phonemes. Vihman (1995) argued that a phonology for word production emerges once the child has around 50 words in a lexicon and as result of an analysis of contrastive sounds that exist across these words. Further, Vihman (1995) proposed children begin with phonetic templates that they fill in as their phonology develops.

With native signers a similar difficulty in identifying first signs arises, but here we have to deal with the dis/continuity of early hand movements and signs (Petitto, 1987; Volterra and Erting, 1994). Before children first sign they babble with the hands (hand open/closes and palm turns), use ritualized gestures and use pointing. Petitto and Marentette (1991) argued that manual babbling changed over time following a process akin to the transition from variegated to canonical babbling in vocal development. In sign, the movement part of the babbling took on a different cadence as it became part of ASL. Cheek et al. (2001) showed that the earliest parameters appearing in sign babbling were those that would appear first in the child's initial sign vocabulary. The analog to this in spoken language acquisition would be where children begin with the simplest and most unmarked speech sounds and then gradually extend their repertoires (Vihman, 1995). For example the phoneme /d/ is one of the first speech sounds to be used systematically during the babbling stage and many children acquiring English produce words with these sounds early, i.e., "dada."

Vihman (1995) proposed that children have an articulatory filter that "screens in" or finds words that are within the child's phonetic capacity and this means the child may understand/perceive a word but avoid producing it, if it included a speech sound that they cannot yet produce. Ideas of selection and filtering can be traced back to the psycholinguistic models of Smith (1973) where a set of innate and universal processes influenced phonological development. As children are developing their phonological representations the acquisition mechanism implements a set of processes which simplify the sound system and account for the "error patterns" that occur in early language. Smith (1973) described three main types of errors: structural "deletion"; "assimilation" processes and finally systemic "substitution." Because these processes were labeled universals they can be tested in sign language acquisition (Clibbens and Harris, 1993; Morgan, 2006).

The rest of this section describes how systemic substitution during the acquisition of sign language feeds into the development of the first level of combinatorial structure in signing: phonology. In speech development substitution means the child replaces the adult target sounds not yet mastered, with sounds already part of their productive speech for example producing "tea" instead of "key" and tar instead of "car." This process is called "fronting" and occurs in typically developing speech until around 4 years. Substitution in children's first words is linked to the markedness of features (complexity and frequency), as well as the child's own small rule system at that point in development.

It is also important to note that the earliest phonological forms (handshapes or vocalizations) are the easiest ones for the child to produce motorically. In many sign language acquisition studies with native signers the handshape component is the most difficult element to articulate correctly and substitution is very common (Boyes Braem, 1990; Clibbens and Harris, 1993; Marentette and Mayberry, 2000; Morgan et al., 2007).

Clibbens and Harris (1993) reported that the child in their study who was acquiring BSL, used only the A (fist) and five (spread fingers) handshapes until 1;7, after which she added the G (index point). Clibbens and Harris (1993) proposed that the differences that occur between a child's production of a sign and the adult target could follow a similar process as those shown for speech and be guided by markedness.

Boyes Braem (1990) proposed a stage model to predict the types of substitutions or simplifications a child might make when acquiring a sign language. As a basic rule if a wrong handshape was used it would come from the same stage or an earlier stage in the model. For example, a child may substitute the five handshape in stage 1 for the B (flat hand fingers closed) handshape in stage 2 of the model. Meier (2005) confirmed this with a larger group of children who followed these same predictable patterns. When they substituted a handshape for an adult target they invariably used one from the first stage of Boyes Braem's model. These patterns can be related to the idea of systemic processes in Smith's (1973) model.

As a parallel to Vihman's (1995) template proposal for spoken language acquisition, Boyes Braem (1990) observed a small set of unmarked handshapes were used by children at the start of their sign acquisition as the initial building blocks of signing phonology. Marentette and Mayberry (2000) and Meier (2005) labeled this first set of handshape phonological "primes." Primes are invariably unmarked forms, maximally perceptually distinctive (i.e., fully open fingers, fully closed, extended single finger), easiest to produce and occur in high frequency in the adult language. The proposed primes for ASL are the "whole five hand," the "fist," and the "index finger" hand and these have been observed in other early child phonologies in other sign languages (e.g., Clibbens and Harris, 1993). Signing children develop a communicative systems based on signs with these prime handshapes. As their vocabulary grows, they attempt to produce more marked handshapes and through well attested phonological processes the output gets re-configured by the child in a systematic way. The claim is that phonological processes feed into the development of componential structure.

Morgan (2006) identified patterns across sign and speech development leading to organization of the phonology. For spoken language development substitutions are typically based on groups of sounds, identified by features. For example, devoicing is a process that may affect all voiced stops. Processes such as devoicing, velar fronting, consonant harmony, or cluster reduction are all different ways to affect groups of sounds (and at the same time, any of these processes may have the result that [t] replaces [d] in a particular instance).

Simplification processes in sign language acquisition are where different primes stand in for visually similar marked handshapes. The idea is that substitutions are through "families" of similar handshapes and far from random. For example the G handshape (index finger) appears as a replacement in all the handshapes that have a "pointing" feature (I, Y, H, F) while not at all as a replacement for the more "fist-like" handshapes (C, W, O, claw 5) where LAX 5 was common (for stills of handshapes see http://en.wiktionary.org/wiki/Appendix:Sign_language_handshapes).

Thus markedness is dealt with in a similar way by the child at the start of language development across modalities. At the same time there are some modality specific features of sign development. One of these concerns the role of the child's own visual feedback. Children acquiring spoken languages get full access to auditory feedback of their own voices. However, because many sign locations are not in the signers visual field (e.g., a sign on the signer's own head) in some cases the child has to produce a sign with less complete feedback. There has not been a lot of research on the role of visual feedback of one's own signs (Emmorey et al., 2009). However, the developmental data suggests it is useful. Young signers make more self-articulation errors with handshapes that are made at locations in peripheral compared with central vision (Ortega and Morgan, 2010).

A second feature of signing that differs to speech is the size of the major articulators. Young children's gross movement development influences there articulation of signs. Two characteristics noted in the literature (Meier, 2005) are proximalization (where distal joint articulation in signs changes to joints closer to the body) and sympathy (one handed signs get changed into two handed ones). There do not appear to be comparable phenomena in spoken development.

INTERIM SUMMARY ON THE ACQUISITION OF PHONOLOGY
Signed and spoken language acquisition is comparable in several ways, with the main overlap for the focus of this review being on how children build componential structure for signs and words. Before children have an established lexicon they use communicative vocal and manual gestures without analyzing sub-lexical contrasts and regularities (Vihman, 1995). Through pressure from a filter/selection model a system of contrasts emerges and in one explanation systemic substitution leads to regularization (Smith, 1973). In signing, the child might look for visual regularities between families of handshapes across the emerging phonology but in both modalities there is an effect of markedness. Children's sensori-motor limitations lead to strategies for reducing markedness in production and this possibly influences

connections between parts of the grammar and the growth of phonological representations (see Newport, 1990 and the "less is more" analogy). Further work is required to test this hypothesis.

In the next section, a mechanism for deriving componential structure is proposed for signing children's acquisition of spatial classifier constructions following the notion of morphological productivity in linguistic development.

COMPARISONS OF THE ACQUISITION OF SPATIAL LANGUAGE IN SIGN AND SPEECH

Many psycholinguistic studies make an assumption that there is enough equivalence between the sign and speech modalities to test out theories of language structure and processing. At the same time there do exist aspects that are more divergent. While cross-linguistically there is a very wide range of ways to talk about space and movement, no spoken language articulates words in an actual space like signed languages do with the classifier system. An English sentence such as "the pen is on the table" encodes the semantic components of figure, ground and location in an arbitrary and language specific way (Talmy, 2003). When signers talk about space they use "classifier" constructions whereby the morphological units of the construction can encode entity and spatial semantics simultaneously in real space (see collections in Emmorey, 2003; Morgan and Woll, 2007).

One linguistic description of these constructions proposes a classifier "template" which carries each semantic component attached to each other in a poly-componential verb (Cogill-Koez, 2000). The figure part of the template is the handshape, the path and or motion is shown by the movement of the hand or relative location. Other information can be fitted into the template such as manner, orientation and simultaneity. The convention in BSL and other signed languages is for the ground to be mentioned first, e.g., the sign TABLE is signed in space in front of the signer by moving two flat hands apart at waist height to create a representation of a surface. Then the template gets filled in: figures are encoded using handshapes that represent classes of referents with shared semantic and visual features (e.g., vehicles or long thin objects). The interesting aspect of signed language classifiers is that they use space to talk about space (Emmorey, 2003). This is a device unlike any other in spoken language but does resemble how hearing people use gestures (Schembri et al., 2005; Marshall and Morgan, in press). Originally classifiers were analyzed as poly-morphemic (Klima and Bellugi, 1979; Supalla, 1986) however, recently there is much debate as to how they might incorporate gesture and as such language acquisition data become relevant.

One intriguing comparison concerns signing children's mastery of the classifier system compared with hearing children. The visual modality might seem much more iconic than words are and would influence the rate and patterning of language development. Here we explore this question using the same two topics described previously: (1) the relationship between gestures and signs and (2) how the child develops combinatorial structure.

In general, spatial language, because of its arbitrariness and cross-linguistic variation is notoriously difficult for children to acquire in spoken languages (Clark, 2004). Although learning

to talk about space in spoken languages begins early it continues for several years. Path expressions emerge in the one- and two-word speech of children in different types of languages. Choi and Bowerman (1991) reported that 14–21-month olds who are learning English produce "out," "up," and "down" to encode their own paths and "on," "in," and "off" for those of objects. By 2 years of age, children use prepositions for encoding topological arrangement of objects, e.g., "on," "above," or "below" (Clark, 2004). Projective relations (e.g., behind) are expressed later: in English, Italian, Serbo-Croatian, and Turkish children do not produce "front/back" (e.g., "the ball is in front of the tree") until about 5 years of age. The use of "left" and "right" to specify the location of one object with respect to another using three-dimensional Euclidean principles appears still later, at about 11 or 12 years (Johnston, 1984). As with other types of morphology the acquisition of spatial language is indexed by productive control over a system (Gentner and Bowerman, 2009). Productivity refers to the acquisition or control of generalizable facts about the system, rather than individualized structures. A child that uses the word "eated" is starting to grasp that the past tense morpheme "ed" can be applied across a class of words in a generalized or productive way (Brown, 1973).

The first studies of the acquisition of classifiers in ASL adopted a poly-morphemic approach and supported this long developmental pattern across modalities (e.g., Schick, 1990). If classifiers are poly-morphemic, children have to grasp the potential to combine semantic contrasts (morphemes encoding an entity "move down," "turn around," "be located next to" etc.) across a system of morphological forms (person, vehicle, flat surface etc). It is not a characteristically successful outcome of language acquisition if children remain with only knowledge of isolated constructions. Control of productive knowledge is far more efficient and offers more expressive power (Brown, 1973; Bybee, 2006).

Slobin et al. (2003) reported early use of classifier handshapes and path descriptions in children learning ASL and Sign Language of the Netherlands (SLN). Slobin et al. (2003), describe a deaf child aged 2;8 with non-native SLN input move a fist with thumb and pinkie extended in a downward arc to express the notion "the plane flies down." Another child exposed to SLN at 2;6 produced two curved spread fingers handshapes and moved them in an upward, slow, zigzag path to show a "balloon drifting away." An even younger child, at 2;1 exposed to ASL, producing a two handed construction where the less-dominant hand, acts as a ground (representing a chair) with a relaxed spread fingers handshape and the dominant hand with the index and middle finger touching and extended, was placed on top the non-dominant hand to encode the figure-ground meaning "the doll stood on a toy chair."

Thus the beginning of the grammar might emerge early and even be available to children who are learning a signed language in less than optimal conditions. Slobin et al. (2003) argued these constructions were precocious compared with hearing children acquiring spoken languages and this was because iconicity and gesture gave the child semantic scaffolds which they later develop into a more formal system. An important issue therefore is how early forms linked to iconicity and gesture get put together in a conventionalized and combinatorial way that corresponds to how

adult native signers use classifiers in a systematic grammar (de Beuzeville, 2006; Slobin, 2008).

The relevant question is at what point does the child have productive control over combinations of meanings and forms rather than just for individual classifiers (Brown, 1973). In his work on the development of first verb constructions in English Tomasello (1992) proposed children begin to use rules for marking semantic contrast item by item. The verb "island" approach describes children applying rules piece-meal before applying them across constructions (Tomasello, 1992).

Brown (1973) established criterion for attributing productive knowledge in a corpus of utterances where forms are analyzed across different tokens and contexts, e.g., "I walked," "Teddy talked," and "Daddy eated," rather than in isolated examples. By looking for productivity in this way we can more easily start to examine how signed language acquisition becomes a conventionalized system of combinations and how this compares with spoken language development.

Following the verb islands concept a signing child hypothetically might use a classifier handshape for a person in only one context, e.g., FATHER CL-PERSON-GO "daddy goes" and not for any other spatial meanings. Later in development the same handshape CL-PERSON could be combined with a different movement or locative expression to describe a person turning, moving next to, over etc. In this way we could see that the verb islands begin to join up and combinatorial structure is emerging: the handshape begins to be combined with other forms to mark more diverse semantic contrasts and is more productive rather than individuated. Morgan et al. (2008) using Brown's productivity criterion described classifier forms in the signing of a native signer between 1;6 and 3;0 and how at the start of his language development gestures were used to describe spatial concepts before the classifier system took over.

COMPONENTIAL STRUCTURE IS DRIVEN BY PRODUCTIVITY

Morgan et al. (2008) described this under-specified use of the classifier system in a case study of native BSL acquisition. They identified gestures, signs and classifiers before looking at how the classifiers developed. Initially they described a usage pattern of sole gestures then gestures combined with parts of the classifier construction template and finally classifiers without gestures. The order of development for spatial language in BSL was:

> whole body as the figure > hand as a the figure and real object as ground or *vice versa* > finger tracing the path > conventional classifier construction.

Between 1;10 and 2;6 there were a set of eight meaningful handshapes that the child used in individual utterances or verb islands. They were not being combined with more than one movement/location component and were thus categorized as non-productive forms. During the age 2;6–3;6 the child expanded the number of different handshape and path/location combinations moving from verb islands to a more productive system. Example, at 2;6 the flat hand was used with three different spatial meanings as was the pinkie/thumb handshape. By 3;0 the two finger handshape was used in three different contexts and the same movement/location components of the classifier template were being

combined with several different handshapes. Thus different parts of the template were being interchanged which suggests the child has more control over the system (see Morgan et al., 2008 for more details).

It is still a debate as to which mechanisms drive productivity in language acquisition: domain general cognitive mechanisms or language-dedicated rules (Tomasello, 1992; Valian, 2009). For the acquisition of the classier constructions once the child has combinatorial structure they can use the system in a productive way. Structure allows productivity and productivity extends combinatorial structure.

SUMMARY ON THE ACQUISITION OF SPATIAL LANGUAGE FROM GESTURES TO SIGNS

The development of verb morphemes in spoken languages typically begins on familiar verbs and repetitive contexts before being used with novel items (Tomasello, 1992). The acquisition of classifiers in signing children thus follows a well attested pattern and so productivity is achieved slowly even with available gestures and iconicity. By waiting for productivity we see that the classifier template gets filled in piecemeal. Productivity is signaled when meaning components start to be interchangeable in the template. While spatial language is used in very different ways in signing and speaking children, this developmental path to a combinatorial structure is familiar and predictable.

LESSONS FROM CHILD SIGN LANGUAGE LEARNERS FOR GENERAL LANGUAGE ACQUISITION THEORY

This paper has presented language acquisition data which documents how mechanisms lead to combinatorial structure in the phonology and morphology of signing. This level of linguistic organization distinguishes signs and gestures. Although they may well be continuous on a spectrum (Kendon, 2004) the acquisition data show that as linguistic systems sign languages are nevertheless subject to typical developmental processes. They are not acquired in a radically different ways to spoken languages but instead conform to how we expect a representational grammar is learned by a child approximating the input and building a conventionalized system.

Findings coming from signing children can inform the general field of language acquisition by firstly emphasizing that linguistic development occurs in universal ways meaning theories that are modality free are preferred. Two more specific observations are also warranted here. We see that native signers use substitution to deal with marked forms by building links between visually related sets of handshapes in their repertoire. These phonological processes can explain why the large group of deaf children who learn sign language late (as they interact with hearing non-signing parents) end up with a different set of phonological abilities when they are adults (Mayberry, 2010). It might be that early reorganization at the sub-lexical level, as a result of maturational limitations, leads children to reap the reward later with more complex phonological representations (Newport, 1990). Secondly, data on spatial language acquisition by signers highlights that children use both holistic gestures and isolated signs initially before arriving at a coherent system with productive linguistic representations. Hearing children acquiring spoken language might also take advantage of the semiotic power of gesture. Early gestures might provide some structure for hearing children to explore meaning and form mappings during language development before their spoken words become part of a productive system. With this in mind continued attention in longitudinal studies of early spoken language development to speech and gesture combinations is worthwhile.

Although native signers are a small number of children their path to the development of componential structure reveals both what is particular about sign language (compared with other visual forms of communication) and what is universal about language acquisition.

REFERENCES

Baker, A., and Woll, B. (eds). (2009). *Sign Language Acquisition.* Amsterdam: John Benjamins. doi: 10.1075/bct.14

Boyes Braem, P. (1990). "Acquisition of the handshape in American Sign Language: a preliminary analysis," in *From Gesture to Language in Hearing and Deaf Children*, eds V. Volterra and C. J. Erting (Washington, DC: Gallaudet University Press), 107–127.

Brentari, D. (1998). *A Prosodic Model of Sign Language Phonology.* Cambridge, MA: MIT Press.

Brentari, D., Coppola, M., Mazzoni, L., and Goldin-Meadow, S. (2012). When does a system become phonological? Handshape production in gesturers, signers, and homesigners. *Nat. Lang. Linguist. Theory* 30, 1–31. doi: 10.1007/s11049-011-9145-1

Brown, R. (1973). *A First Language: The Early Stages.* Cambridge, MA: Harvard University Press. doi: 10.4159/harvard.9780674732469

Bybee, J. L. (2006). From usage to grammar: the mind's response to repetition. *Language* 82, 711–733. doi: 10.1353/lan.2006.0186

Chamberlain, C., Morford, J. P., and Mayberry, R. I. (2000). *Language Acquisition by Eye.* Mahwah, NJ: Lawrence Erlbaum Associates.

Cheek, A., Cormier, K., Repp, A., and Meier, R. P. (2001). Prelinguistic gesture predicts mastery and error in the production of early signs. *Language* 77, 292–323. doi: 10.1353/lan.2001.0072

Chen Pichler, D. (2012) "Language acquisition," in *Handbook of Linguistics and Communication Science: Sign Language*, eds R. Pfau, B. Woll, and M. Steinbach (Berlin: de Gruyter), 647–686.

Choi, S., and Bowerman, M. (1991). Learning to express motion events in English and Korean: the influence of language-specific lexicalization patterns. *Cognition* 41, 83–121. doi: 10.1016/0010-0277(91)90033-Z

Chomsky, N. (1965). *Aspects of the Theory of Syntax.* Cambridge, MA: MIT Press.

Clark, E. (2004). How language acquisition builds on cognitive development. *Trends Cogn. Sci.* 8, 472–478. doi: 10.1016/j.tics.2004.08.012

Clibbens, J., and Harris, M. (1993). "The acquisition of formational parameters in British Sign Language: a case study," in *Critical Aspects of Language Acquisition and Development*, eds D. Messer and G. Turner (London: Macmillan), 197–208.

Cogill-Koez, D. (2000). A model of signed language 'classifier predicates' as templated visual representations. *Sign Lang. Linguist.* 3, 209–236. doi: 10.1075/sll.3.2.04cog

de Beuzeville, L. (2006). *Visual and Linguistic Representation in the Acquisition of Depicting Verbs: A Study of Native Signing Deaf Children of Auslan (Australian Sign Language).* Ph.D. thesis, University of Sydney, Sydney.

de Graff, M. (1999). "Creolization, language change and language acquisition: an epilogue," in *Language Creation and Language Change: Creolization, Diachrony, and Development*, ed. M. DeGraff (Cambridge, MA: MIT Press), 473–543.

de Ruiter, J. (2000). "The production of gesture and speech," in *Language and Gesture*, ed. D. McNeill (Cambridge: Cambridge University Press), 284–311.

Duncan, S. (2003). "Gesture in language: issues for sign language research," in *Perspectives on Classifier Constructions*, ed. K. Emmorey (Mahwah, NJ: Lawrence Erlbaum Associates), 259–268.

Emmorey, K. (2002). *Language, Cognition, and the Brain: Insights from Sign Language Research.* Mahwah, NJ: Lawrence Erlbaum and Associates.

Emmorey, K. (ed.). (2003). *Perspectives on Classifier Constructions in Sign Languages.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Emmorey, K., Gertsberg, N., Korpics, F., and Wright, C. E. (2009). The influence of visual feedback and register changes on sign language production: a kinematic study with deaf signers. *Appl. Psycholinguist.* 30, 187–203. doi: 10.1017/S0142716408090085

Gentner, D., and Bowerman, M. (2009). "Why some spatial semantic categories are harder to learn than others: the typological prevalence hypothesis," in *Crosslinguistic Approaches to the Psychology of Language: Research in the Tradition of Dan Isaac Slobin*, eds J. Guo, E. Lieven, N. Budwig, S. Ervin-Tripp, K. Nakamura, and S. Ozcaliskan (New York: Psychology Press), 465–480.

Goldin-Meadow, S. (2003). *The Resilience of Language: What Gesture Creation in Deaf Children Can Tell Us About How All Children Learn Language.* New York: Psychology Press.

Goldin-Meadow, S., McNeill, D., and Singleton, J. (1996). Silence is liberating: removing the handcuffs on grammatical expression in the manual modality. *Psychol. Rev.* 103, 34–55. doi: 10.1037/0033-295X.103.1.34

Goldin-Meadow, S., Mylander, C., and Butcher, C. (1995). The resilience of combinatorial structure at the word level: morphology in self-styled gesture systems. *Cognition* 56, 195–262. doi: 10.1016/0010-0277(95)00662-I

Johnston, J. R. (1984). The acquisition of locative meanings: behind and in front of. *J. Child Lang.* 2, 407–422.

Jusczyk, P. W., Smolensky, P. and Allocco, T. (2002). How English-learning infants respond to markedness and faithfulness constraints. *Lang. Acquis.* 10, 31–73. doi: 10.1207/S15327817LA1001_3

Kendon, A. (2004). *Gesture: Visible Action as Utterance.* Cambridge: Cambridge University Press.

Klima, E., and Bellugi, U. (1979). *The Signs of Language.* Cambridge, MA: Harvard University Press.

Liddell, S. (2003). *Grammar, Gesture and Meaning in American Sign Language.* Cambridge: Cambridge University Press. doi: 10.1017/CBO9780511615054

Marentette, P. F., and Mayberry, R. I. (2000). "Principles for an emerging phonological system: a case study of early ASL acquisition," in *Language Acquisition by Eye*, eds C. Chamberlain, J. Morford, and R. Mayberry (Mahwah, NJ: Erlbaum), 71–90.

Marshall, C., and Morgan, G. (in press). From gesture to sign language: conventionalisation of classifier constructions by adult hearing learners of British Sign Language. *Topics Cogn. Sci.*

Mayberry, R. I. (2010). "Early language acquisition and adult language ability: what sign language reveals about the critical period for language," in *Oxford Handbook of Deaf Studies, Language, and Education*, Vol. 2, eds M. Marschark and P. Spencer (New York: Oxford University Press), 281–291.

McNeill, D. (1992). *Hand and Mind.* Chicago: University of Chicago Press.

McNeill, D., and Duncan, S. D. (2000). "Growth points in thinking-for-speaking," in *Language and Gesture*, ed. D. McNeill (Cambridge: Cambridge University Press), 141–161.

Meier, R. P. (2005). "The forms of early signs: explaining signing children's articulatory development," in *Advances in the Sign-Language Development of Deaf and Hard-of-Hearing Children*, eds B. Schick, M. Marschark, and P. Spencer (Oxford: Oxford University Press), 202–230.

Mitchell, R. E., and Karchmer, M. A. (2004). Chasing the mythical ten percent: parental hearing status of deaf and hard of hearing students in the United States. *Sign Lang. Stud.* 4, 138–163. doi: 10.1353/sls.2004.0005

Morgan, G. (2006). Children are just lingual: the development of phonology in British Sign Language (BSL). *Lingua* 116, 1507–1523. doi: 10.1016/j.lingua.2005.07.010

Morgan, G., Barrett-Jones, S., and Stoneham, H. (2007). The first signs of language: phonological development in British Sign Language. *Appl. Psycholinguist.* 28, 3–22. doi: 10.1017/S0142716407070014

Morgan, G., Herman, R., Barriere, I., and Woll, B. (2008). The onset and mastery of spatial language in children acquiring British Sign Language. *Cogn. Dev.* 23, 1–9. doi: 10.1016/j.cogdev.2007.09.003

Morgan, G., and Woll, B. (eds.) (2002). *Directions in Sign Language Acquisition.* Amsterdam: Benjamins. doi: 10.1075/tilar.2

Morgan, G. and Woll, B. (eds.) (2007). Understanding sign language classifiers through a polycomponential approach. *Lingua* 117, 1159–1168. doi: 10.1016/j.lingua.2006.01.006

Newport, E. L. (1990). Maturational constraints on language learning. *Cogn. Sci.* 14, 11–28. doi: 10.1207/s15516709cog1401_2

Ortega, G., and Morgan, G. (2010). Comparing child and adult development of a visual phonological system. *Lang. Interact. Acquis.* 1, 67–81. doi: 10.1075/lia.1.1.05ort

Petitto, L. A. (1987). On the autonomy of language and gesture: evidence from the acquisition of personal pronouns in American Sign Language. *Cognition* 27, 1–52. doi: 10.1016/0010-0277(87)90034-5

Petitto, L. A. (1997). "In the beginning: on the genetic and environmental factors that make early language acquisition possible," in *The Inheritance and Innateness of Grammars*, ed. M. Gopnik (Oxford: Oxford University Press), 45–69.

Petitto, L. A., and Marentette, P. (1991). Babbling in the manual mode: evidence for the ontogeny of language. *Science* 251, 1483–1496. doi: 10.1126/science.2006424

Pfau, R., Steinbach, M., and Woll, B. (eds.) (2012). *Sign Language: An International Handbook.* Boston, MA: de Gruyter. doi: 10.1515/9783110261325

Sandler, W., and Lillo-Martin, D. (2006). *Sign Language and Linguistic Universals.* Cambridge: Cambridge University Press. doi: 10.1017/CBO9781139163910

Sandler, W., Meir, I., Padden, C., and Aronoff, M. (2005). The emergence of grammar: systematic structure in a new language. *Proc. Natl. Acad. Sci. U.S.A.* 102, 2661–2665. doi: 10.1073/pnas.0405448102

Schembri, A., Jones, C., and Burnham, D. (2005). Comparing action gestures and classifier verbs of motion: evidence from Australian Sign Language, Taiwan Sign Language, and non-signers' gestures without speech. *J. Deaf Stud. Deaf Educ.* 10, 272–290. doi: 10.1093/deafed/eni029

Schick, B. (1990). Classifier predicates in American Sign Language. *Int. J. Sign Linguist.* 1, 15–40.

Schick, B. (2004). "Acquiring a visually-motivated language: evidence from diverse learners," in *Advances in the Sign Language Development of Deaf and Hard-of-Hearing Children*, eds B. Schick, M. Marschark, and P. Spencer (New York: Oxford University Press).

Senghas, A., Kita, S., and Ozyurek, A. (2004). Children creating core properties of language: evidence from an emerging sign language in Nicaragua. *Science* 305, 1779–1782. doi: 10.1126/science.1100199

Singleton, J., and Newport, E. (2004). When learners surpass their models: the acquisition of American Sign Language from inconsistent input. *Cogn. Psychol.* 49, 370–407. doi: 10.1016/j.cogpsych.2004.05.001

Singleton, J. L., Goldin-Meadow, S., and McNeill, D. (1995). "The cataclysmic break between gesticulation and sign: evidence against a unified continuum of gestural communication," in *Language, Gesture, and Space*, eds K. Emmorey and J. Reilly (Hillsdale, NJ: Erlbaum Associates), 287–311.

Slobin, D. I. (2008). Putting the pieces together: commentary on "The onset and mastery of spatial language in children acquiring British Sign Language" by G. Morgan, R. Herman, I. Barriere, and B. Woll. *Cogn. Dev.* 23, 20–23. doi: 10.1016/j.cogdev.2007.09.001

Slobin, D. I., Hoiting, N., Kuntze, K., Lindert, R., Weinberg, A., and Pyers, J. (2003). "A cognitive/functional perspective on the acquisition of "classifiers," in *Perspectives on Classifier Constructions in Sign Languages*, ed. K. Emmorey (Mahwah, NJ: Lawrence Erlbaum Associates), 271–296.

Smith, N. V. (1973). *The Acquisition of Phonology.* Cambridge: Cambridge University Press.

Stokoe, W. (1960). Sign language structure: an outline of the visual communication systems of the American deaf. *Stud. Linguist. Occas. Pap.* 8, 1–78.

Supalla, T. (1986). "The classifier system in American sign language," in *Noun Classes and Categorization*, ed. C. G. Craig (Amsterdam: John Benjamins), 181–214. doi: 10.1075/tsl.7.13sup

Talmy, L. (2003). "The representation of spatial structure in spoken and signed language," in *Perspectives on Classifier Constructions in Sign Languages*, ed. K. Emmorey (Mahwah, NJ: Lawrence Erlbaum Associates), 271–296.

Tomasello, M. (1992). *First Verbs.* Cambridge: Cambridge University Press.

Valian, V. (2009). "Innateness and learnability," in *Handbook of Child Language*, ed. E. Bavin (Cambridge: Cambridge University Press), 15–34. doi: 10.1017/CBO9780511576164.002

Van Loon, E., Pfau, R., and Steinbach, M. (in press). "The grammaticalization of gestures in sign languages," in *Body - Language - Communication: An International Handbook on Multimodalityl in Human Interaction*, eds C. Müller, A. Cienki, E. Fricke, S. H. Ladewig, D. McNeill, and S. Tessendorf (Berlin: Mouton de Gruyter).

Vihman, M. (1995). *Phonological Development: Origins of Language in the Child.* Oxford: Blackwell.

Volterra, V., and Erting, C. (eds). (1994). *From Gesture to Language in Hearing and Deaf Children.* Washington, DC: Gallaudet University Press.

Zeshan, U. (2003). "Classificatory constructions in Indo-Pakistani Sign Language: grammaticalization and lexicalization process," in *Perspectives on Classifier Constructions in Sign Languages,* ed. K. Emmorey (Mahwah, NJ: Psychology Press), 41–113.

Conflict of Interest Statement: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Language choice in bimodal bilingual development

Diane Lillo-Martin[1,2]*, Ronice M. de Quadros[3], Deborah Chen Pichler[4] and Zoe Fieldsteel[5]

[1] Department of Linguistics, University of Connecticut, Storrs, CT, USA
[2] Haskins Laboratories, New Haven, CT, USA
[3] Departamento de Libras, Universidade Federal de Santa Catarina, Florianópolis, Brazil
[4] Department of Linguistics, Gallaudet University, Washington, DC, USA
[5] Department of Linguistics, Brown University, Providence, RI, USA

Bilingual children develop sensitivity to the language used by their interlocutors at an early age, reflected in differential use of each language by the child depending on their interlocutor. Factors such as discourse context and relative language dominance in the community may mediate the degree of language differentiation in preschool age children. Bimodal bilingual children, acquiring both a sign language and a spoken language, have an even more complex situation. Their Deaf parents vary considerably in access to the spoken language. Furthermore, in addition to code-mixing and code-switching, they use code-blending—expressions in both speech and sign simultaneously—an option uniquely available to bimodal bilinguals. Code-blending is analogous to code-switching sociolinguistically, but is also a way to communicate without suppressing one language. For adult bimodal bilinguals, complete suppression of the non-selected language is cognitively demanding. We expect that bimodal bilingual children also find suppression difficult, and use blending rather than suppression in some contexts. We also expect relative community language dominance to be a factor in children's language choices. This study analyzes longitudinal spontaneous production data from four bimodal bilingual children and their Deaf and hearing interlocutors. Even at the earliest observations, the children produced more signed utterances with Deaf interlocutors and more speech with hearing interlocutors. However, while three of the four children produced >75% speech alone in speech target sessions, they produced <25% sign alone in sign target sessions. All four produced bimodal utterances in both, but more frequently in the sign sessions, potentially because they find suppression of the dominant language more difficult. Our results indicate that these children are sensitive to the language used by their interlocutors, while showing considerable influence from the dominant community language.

Keywords: bimodal bilingualism, bilingual development, code-blending, language mixing, interlocutor sensitivity

INTRODUCTION

There has been much interest in how the languages of children developing as simultaneous bilinguals separate and interact. It has frequently been observed that, especially at the earliest ages, children may seem to mix their languages, by using structures that apparently combine elements of both (Grosjean, 1989; Bhatia and Ritchie, 1999; Paradis, 2001). In addition, children may interact with speakers of one of their languages (say, language A) using elements of their other language (say, language α)—showing incomplete discourse separation (Paradis and Nicoladis, 2007). Such observations have led to the proposal that bilingual children's language is "fused" at an early age (Volterra and Taeschner, 1978); that is, they have one grammar with elements of both languages.

However, many authors have argued against the view that bilingual children's languages are "fused" (Genesee, 1989). They observe, for example, that even highly fluent bilingual adults produce "mixed" structures showing elements of both languages.

Adult bilinguals who are fully proficient in both languages allow the languages to interact in varied and interesting ways (Costa and Santesteban, 2004; Bishop and Hicks, 2005; Gonzalez-Vilabazo and López, 2012). Code-switching is taken as a sign of bilingual proficiency (Poplack, 1980; Lucas and Valli, 1992), and it is heavily used as an in-group sociolinguistic phenomenon in highly bilingual communities (Bhatt and Bolonyai, 2011). Nevertheless, it cannot be said that young bilingual children's languages are completely separate (Unsworth, 2013). We conclude, then, that the best tack to take toward understanding the development of bilingualism is to model the adult state and to see how children move toward achieving this state.

Our project takes this approach: we are developing a model of bilingualism that we expect applies equally to describing both adult and child states, although some of the details of grammatical knowledge for children may be different from that of adults. Our project also takes one further step: while it should also apply to unimodal bilingualism, we are developing this model

in the context of bimodal bilingualism: children who are becoming bilingual in a signed language and a spoken language (for a general overview on such children, see Baker and van den Bogaerde, 2014). Bimodal bilinguals can be hearing (using the spoken and written form of a spoken language) or Deaf (some using both forms, others using only the written form of a spoken language). They include people who use a sign language casually, daily, or professionally as interpreters. Most of the children we are studying—and all of the ones in the current report—have normal hearing, but their families (in particular, one or both parents) are Deaf and use a sign language with them. The children acquire sign language at home, and they acquire spoken language from the greater community (including other relatives, neighbors, schools, etc.). Then, we ask how the issues around language separation and mixing are different in the context of bimodal bilingualism.

The few existing studies with adult bimodal bilinguals have led to several conclusions. First, as with unimodal bilinguals, both of the languages of bimodal bilinguals are active and influence language use and processing, even in contexts that only call for one language (Kroll and Stewart, 1994; Kroll et al., 2006; Emmorey et al., 2008b; Shook and Marian, 2012). The various types of language mixing observed in unimodal bilinguals can also be found, but with a twist. Code-switching—in this context, ceasing production in one language (e.g., speech) and starting up in the other language (e.g., sign)—is relatively rare. Emmorey et al. (2008a) studied adult bimodal bilinguals, often known as codas ("Child of Deaf adult," implied hearing and adult), in a highly bilingual context (conversing with another familiar coda). Overall, their participants produced code-switches in only 6.26% of the utterances analyzed. However, they displayed another type of language "mixing," unique to bimodal bilinguals: code-blending. Code-blending is the natural and spontaneous use of speech and sign together. In the data collected by Emmorey et al. (2008a) 35.71% of all utterances contained code-blending. Finally, Emmorey et al. also observed the use of sign language structures in the spoken language—so-called cross-linguistic influence, or transfer—another type of language "mixing."

Bimodal bilinguals introduce a new type of "mixing" to the picture of how the languages of a bilingual interact. Not only do they produce structures showing cross-linguistic influence and code-switching, they also productively use code-blending. Any model of bilingualism—the target toward which children develop—must account for all three of these phenomena.

In a series of works, we have been developing such a model (Lillo-Martin et al., 2010, 2012; Koulidobrova, 2012; Quadros et al., 2013). Our model, illustrated in **Figure 1**, adopts the viewpoint that bilingualism should be explained using the same architecture of linguistic behavior as required for monolinguals (MacSwan, 2000, 2005). Bilinguals simply have additional materials to work with, but they must adhere to the overall grammatical possibilities and constraints placed on any language. We start with a standard generative perspective incorporating concepts of distributed morphology (Halle and Marantz, 1993; Chomsky, 1995). The input to a derivation contains abstract roots and morphemes. For a bilingual, there are two sets of items to choose from for every derivation. During the syntax, featural requirements must

FIGURE 1 | Synthesis model (Lillo-Martin et al., 2012; Quadros et al., 2013).
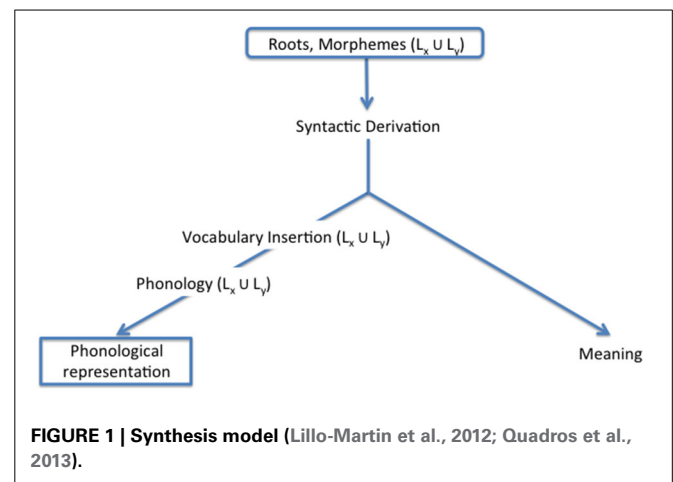
be satisfied; and in some cases, elements from language A may satisfy the requirements of elements from language α, leading to structures with cross-linguistic influence or transfer. At the point of Vocabulary Insertion, elements from either language may be inserted, as long as all featural requirements are satisfied, leading to code-switching. Finally, when two independent sets of articulators are available, lexical items from both languages are possible, making code-blending possible. All three of these outcomes are considered natural consequences of our Synthesis model, so-called because it offers a picture of the combinatorial possibilities allowed by the language architecture.

Our project tests the usefulness of this model in explaining the development of bimodal bilingualism. We have found that hearing children acquiring a sign language and a spoken language (kodas—kids of Deaf adults) engage in the types of productions predicted by the Synthesis model: transfer, code-switching, and code-blending (see references cited in previous paragraph). Given that code-blending is an option available to bimodal bilinguals and not to unimodal bilinguals, we now raise the question whether the process of developing interlocutor sensitivity and discourse separation of languages is different for these two groups of children. How do koda children employ code-blending in their developing language selection? In addition, since parents and other interlocutors vary in their own use of code-blending, how do children adjust to the modality of the input in a given situation?

In this article, we address this question by presenting data from our study on the development of bimodal bilingualism in children learning one of two language pairs: American Sign Language (ASL) and English (Eng) in the US, or Brazilian Sign Language (Libras) and Brazilian Portuguese (BP) in Brazil. The data from two children for each language pair indicate that 2-year-old kodas are sensitive to their interlocutor, and modulate their language choice accordingly, but they are also influenced by the fact that the spoken languages are dominant in their broader community and they do not simply mirror the language choices of their interlocutors. Note that our use of the term "choice" is not meant to necessarily imply a conscious decision; it is simply the term to describe the language used by the child or adult in a particular situation.

BACKGROUND

PREVIOUS STUDIES ON LANGUAGE CHOICE IN THE DEVELOPMENT OF UNIMODAL BILINGUALS

Studies of unimodal bilingual children have found that they typically display interlocutor sensitivity at an early age, using more of language A with an interlocutor who speaks A, and more of language α with an interlocutor who speaks α (Genesee et al., 1995; Petitto et al., 2001). This does not mean that the child will only use A or α with speakers of A or α, respectively, or even mostly A/α in the "appropriate" environment. As Paradis and Nicoladis (2007, p. 278) summarize, "Interlocutor sensitivity, then, is not the same as perfect separation of language by discourse context (discourse separation)."

The child's degree of interlocutor sensitivity changes over the early years. At the earliest ages (before 2;0), children's language choices may be attenuated by their lexical knowledge, since a certain amount of code-switching might take place to fill lexical gaps (Deuchar and Quay, 1999; Nicoladis and Secco, 2000). Deuchar and Quay argued that when lexical knowledge is taken into consideration (i.e., considering whether or not the appropriate language is used when the child knows the word in both languages), "there is a strong tendency for the language of the child's utterances to match that of the context" as early as 1;07–1;08.

Genesee et al. (1995) and Nicoladis and Genesee (1996), and others have observed that 2-year-old bilingual children generally demonstrate interlocutor sensitivity. During this period, there are several factors presumed to contribute to the degree of sensitivity and discourse separation children display. One factor is language dominance: children are more likely to use their dominant language in the contexts calling for it than they are to use their own non-dominant language in its contexts (Genesee et al., 1995; Nicoladis and Genesee, 1996). Another relevant factor is the communication style used in the home. When parents are more tolerant and indicate understanding when their children code-mix or choose the "inappropriate" language (sometimes known as a bilingual strategy), children may display less discourse separation, compared to families who are more strict in their expectations about language choices (that is, they pursue a one-parent one-language or "monolingual" strategy) (Döpke, 1992; Lanza, 1997).

Some studies report a high degree of sensitivity and control over language choice at a relatively early age. Comeau et al. (2003) studied six 2-year-old French-English bilinguals (2;00–2;07; mean 2;05). In their study, an experimenter interacted with the children on three separate occasions, deliberately modifying her rate of code-mixing from 15% of the time in the first session, to 40% in the second session, and back to 15% in the third session. Remarkably, they found that five of the six children matched the changes in proportion of mixing overall, and almost all comparisons showed that the children were more likely to use a mixed utterance following a mixed utterance by the interlocutor, and a non-mixed utterance following a non-mixed utterance. These results demonstrate a very early ability to make language choice selections to match those encouraged by the context.

One study examined the interlocutor sensitivity of slightly older children, in order to determine whether true discourse separation can be achieved in the preschool years. In addition to taking into consideration children's relative language dominance, this study also considered the factor of community dominance. Paradis and Nicoladis (2007) studied eight children, ages 3;06–4;11, in the English-dominant English-French bilingual community of Alberta, Canada. In this broader context, people are more likely to use English-only with English-speaking interlocutors, with some mixing occurring with French-speaking interlocutors. As expected (see **Figure 2**), the French-dominant children in this study tended to use French-only in French contexts, and they were highly likely to use English-only in English contexts. On the other hand, while the English-dominant children used English virtually exclusively in the English contexts, they had a lower proportion use of French in the French contexts. Paradis and Nicoladis suggested that the dominance of English in the greater sociolinguistic context contributed to this result; indeed, there was very little mixing in English contexts. In French contexts, more mixing was tolerated, with the children with weaker skills in French responsible for a good deal of this mixing.

PREVIOUS STUDIES ON CODE-BLENDING AND LANGUAGE CHOICE IN THE DEVELOPMENT OF YOUNG BIMODAL BILINGUALS

Studies of code-blending and language choice for pre-school aged bimodal bilinguals are still fairly rare, although interest in this topic stretches back several decades. All of the previous studies, like ours, focus on kodas—hearing children with at least one Deaf signing parent. Very early investigations included that by Griffith (1985), a longitudinal study of the hearing son of two Deaf American parents, with a Deaf older sibling. Griffith reports that
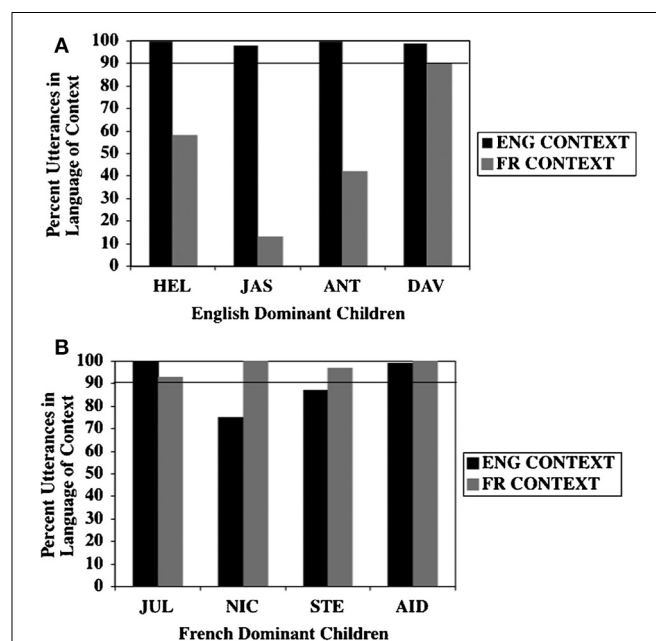

FIGURE 2 | Unimodal bilinguals. Percent use of utterances in the language of the context by **(A)** English-dominant children and **(B)** French-dominant children in English and French contexts (Paradis and Nicoladis, 2007) (Reproduced with permission from Taylor and Francis).

the bimodal bilingual child from the age of 19 months demonstrated "mode-switching" or the use of different language choices according to his interlocutors. Over time, he matched the "mode" most frequently used by each partner, signing more with his sign-dominant father and using sign+speech with his mother, who tended to address him in like manner. Griffith proposed that the child deduced the language preferences of his interlocutors based on whether or not they reacted to his speech only, sign only and sign+speech utterances. Further evidence that the child engaged in such "mode-finding" analysis came from his sessions with new, unfamiliar conversational partners, during which he appeared to try out various conversational modes and watch for the reaction of his interlocutor. Overall, Griffith concluded that her bimodal bilingual subject displayed considerable and early communicative competence in selecting an appropriate communication mode according to his interlocutor(s).

More recent investigations on code-blending reveal a more complicated picture of developing language choice among very young bimodal bilinguals. In a series of reports on their longitudinal, spontaneous production data from three Dutch hearing children and their Deaf mothers, van den Bogaerde and Baker (2005, 2009; also van den Bogaerde, 2000; Baker and van den Bogaerde, 2008) pointed out that language usage patterns do not necessarily remain static, and that language choices of both bimodal bilingual children and their mothers can change over time (see also Kanto et al., 2013). van den Bogaerde and Baker (2009) reported that the mothers in their study all used a high and fairly consistent percentage of code-blended utterances with their children across three sampling times (when the children were aged 1;06, 3;00, and 6;00). All three mothers also increased their use of NGT-only (Sign Language of the Netherlands) production over time. The bimodal bilingual children in the study increased their use of code-blending overall between 1;06 and 6;00 to levels similar to their mothers', but the same was not true for their production of NGT-only utterances. Two of the three children also continued to produce a much greater proportion of spoken Dutch utterances by 6;00 than was present in their mothers' input. These patterns are illustrated in **Figure 3**, showing the production of Dutch, NGT and code-blended utterances over time by the children and their mothers, respectively. Note that van den Bogaerde and Baker did not consider phonation to be a criterion for code-blending. Thus, signed utterances accompanied by mouthing of Dutch words, even in the complete absence of any voicing, were counted as code-blending in their data. While some researchers also adopt this practice (e.g., Fung, 2012, studying code-blending in Hong Kong Sign Language and Cantonese), most others (including us) either explicitly or implicitly consider an utterance to include blending only if sign is accompanied by speech with phonation or at least whispering (e.g., Petitto et al., 2001; Emmorey et al., 2008a; Bishop, 2010; Chen Pichler et al., 2010; Donati and Branchini, 2013; Kanto et al., 2013; Petroj et al., 2014).

Van den Bogaerde and Baker concluded from their data that the language choices of the bimodal bilingual children can only be partially explained by input patterns. Other potential influences, such as the children's language proficiency in NGT vs. spoken Dutch, number of Deaf members in the immediate family, and
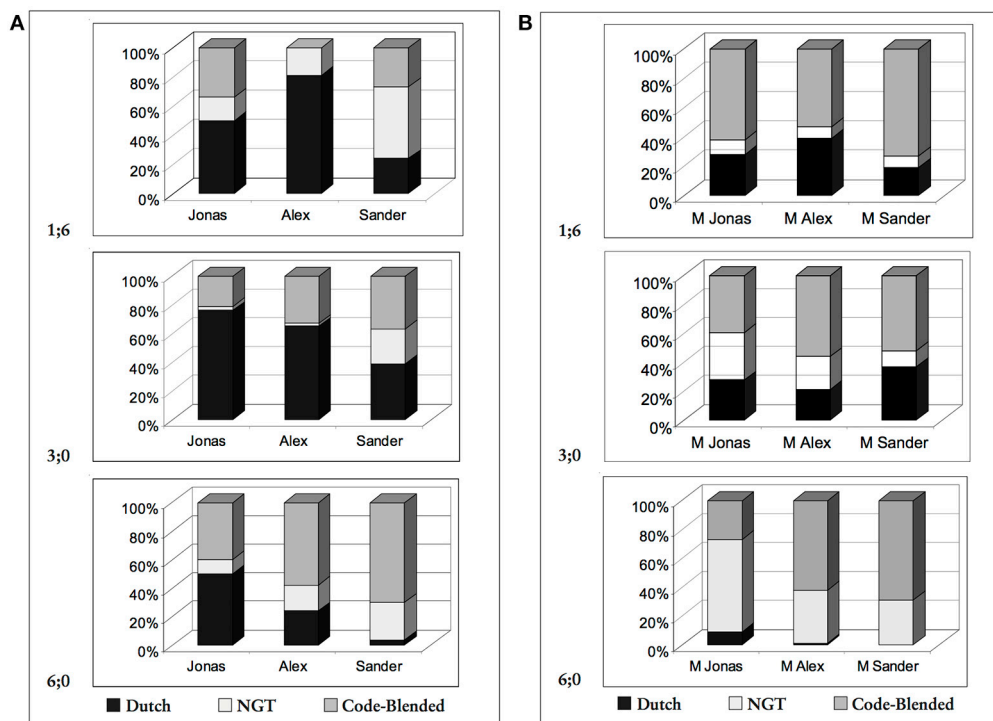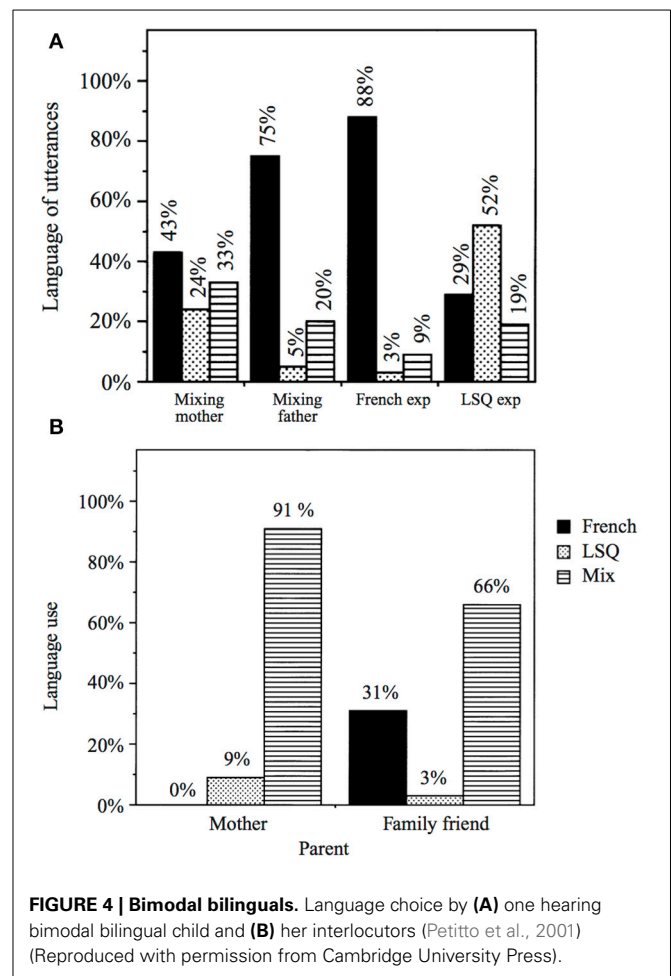

FIGURE 3 | Bimodal bilinguals. Language choice by **(A)** hearing children and **(B)** deaf mothers at ages 1;06, 3;00, and 6;00 (van den Bogaerde and Baker, 2009) (Reproduced with permission from Gallaudet University Press).

changes in language environment (i.e., entry into school, a Dutch-only environment), also exerted only temporary or inconclusive effects on children's code-blending production. On the other hand, the authors observed that the degree to which mothers tolerated being addressed in speech seemed to have an effect on the children's language choices. Support for this idea comes from the one child in the study, Sander, whose language choice over time most closely resembled that of his mother, whom he addressed almost exclusively in NGT or NGT-Dutch blends by 6;00. Van den Bogaerde and Baker noted that Sander's mother often urged him to sign, even when she could understand his speech perfectly well, prompting the authors to propose mothers' choice of a "more monolingual or bilingual strategy" as the best predictor of bimodal bilingual children's language production patterns. This overall conclusion is similar to that discussed earlier with respect to Döpke (1992) and Lanza (1997) studies of unimodal bilinguals.

A similar conclusion was reached by Kanto et al. (2013), who reported that Finnish kodas whose Deaf parents addressed them primarily in sign showed more development in FinSL (Finnish Sign Language) vocabulary and syntactic complexity from 12 to 36 months than their counterparts who were addressed in mixed sign and speech. The former group's sign exposure was also enhanced by regular weekly/biweekly interactions with Deaf individuals besides their parents, although no information was available on the degree to which these other Deaf individuals mixed sign and speech.

Several of the language mixing patterns reported by van den Bogaerde and Baker were also observed by Petitto et al. (2001) for three LSQ (Québec Sign Language)-French bimodal bilingual children and their caretakers. Like the Deaf Dutch mothers, the Deaf caretakers in the Petitto et al. (2001) study also employed a significant degree of code-mixing in their input to their koda children, although the authors did not specify the relative proportions of code-switching vs. code-blending for the parental data. The three children were observed from roughly 0;10–1;08 for the youngest child, 2;10–3;04 for the middle child, and 3;09–4;03 for the oldest child. They were filmed interacting with their Deaf parents, as well as with unfamiliar experimenters who behaved as if they were monolingual in either French or LSQ, allowing observation of the children's reactions to novel communicative environments that called for only spoken language or only sign language. As for their Dutch counterparts, code-blended utterances made up a notable percentage of the utterances these LSQ-French bilinguals addressed to their interlocutors, particularly for the two older children. Petitto et al. (2001) attributed children's degree of mixing directly to the degree of mixing in parental input, citing the relatively high percentages of mixing produced by the second bimodal bilingual subject to her parents (20–33%) and the very high percentage of mixing present in her parents' utterances (66–91%) (see **Figure 4** for this child's results). In contrast, French-English comparison bilinguals in their study whose parents addressed them in only one language or the other produced virtually no mixes at all.

However, like Petitto et al. (2001), van den Bogaerde and Baker (2009) concluded that input patterns alone were not sufficient to predict the language choices of their young bimodal bilingual subjects. They cited early sensitivity to interlocutor language

FIGURE 4 | Bimodal bilinguals. Language choice by **(A)** one hearing bimodal bilingual child and **(B)** her interlocutors (Petitto et al., 2001) (Reproduced with permission from Cambridge University Press).

and the child's own language preference as two additional factors accounting for the children's language choices. The authors argued that children's sensitivity to interlocutor language could be detected despite the fact that inappropriate language choices were still fairly frequent in the children's production data. Crucially, the children modified their relative proportion use of one language or another across interlocutors with different language needs. This pattern was especially evident in the two data collection conditions in which the children interacted with novel experimenters who behaved as if they were monolingual in either French or LSQ. For instance, **Figure 4A** shows that although this child used a considerable amount of LSQ and mixing with her parents, she reduced both of these categories dramatically and increased her use of French-only utterances to 88% while interacting with a novel experimenter who spoke only in French. Such modification of proportions of language use was evident from the youngest children in both the LSQ-French and comparison French-English groups. Petitto et al. (2001) argued that the cases of inappropriate language choice were a developmental feature, most likely due to children's language preference and/or temporary lexical gaps, and did not diminish the evidence for "a clear capacity to alter their language choices depending upon the specific language of the addressees, despite differences in degree" (2001, 479).

Petitto et al. (2001) observed that even in blending, children combined signs and speech in semantically appropriate ways to create a cohesive single proposition. Furthermore, when children occasionally produced equivalent strings of signs and speech in different word orders, they chose word orders appropriate for each language. Petitto et al. interpreted such examples as strong evidence that bimodal language mixing was "systematic and principled" (2001, p. 488) from children's earliest utterances, indicating that they differentiated between their two grammars, and refuting popular concerns of language mixing as a sign of language confusion.

Code-blended utterances produced by young kodas are typically quite short, many of them consisting of a single sign plus a single word. In contrast, older children and adult codas are capable of producing much longer code-blended utterances, resulting in more complex interactions between the speech and sign (Emmorey et al., 2008a; Donati and Branchini, 2013). In our on-going work, we investigate the code-blending produced by the children in our project in more detail.

EXPECTATIONS FOR THE CURRENT STUDY

Taking into consideration the previous studies with unimodal and bimodal bilinguals, the present study was designed to investigate a series of research questions about interlocutor sensitivity, the role of the input, and the unique possibilities for language mixing that emerge in the context of bimodal bilingualism.

In particular, bimodal bilinguals, unlike unimodal bilinguals, have the possibility of three modalities of expression: speech, sign, or bimodal. When addressing various interlocutors, children may take into consideration their ability to understand language addressed to them in each modality. In particular, some Deaf interlocutors have limited access to speech, but this varies greatly from person to person. Some Deaf parents may use speech or blending with their children, or may indicate understanding of spoken or blended utterances addressed to them. Others may insist on sign or blending, which permits the message to be conveyed in sign as well as in speech. Thus, for complete discourse separation, bimodal bilingual children might not be expected to use only sign in sign contexts, but a combination of sign and blending. Furthermore, given the possibility that separation is more complete in the language which is dominant for the community, it might be that as bimodal bilingual children develop, they use more speech-only production in speech contexts, even if a greater variety of choices are made in sign contexts. Note that it is not possible for us to take into consideration children's own language dominance, as there is no independent measure available that is comparable across the sign languages and the spoken languages.

Research Question 1: Do developing bimodal bilingual children show interlocutor sensitivity by selecting language modality at differential rates in speech and sign target sessions? In particular, do they show a greater proportion of spoken language in Speech-target sessions and a greater proportion of sign language in Sign-target sessions? The null hypothesis is that children's language selection does not vary by context (target language). Our expectation is that there will be a difference in language selection across different target language sessions.

Research Question 2: If there is any difference between Speech-target sessions and Sign-target sessions, is this influenced by the dominance of the spoken language in the broader sociolinguistic context? Although the child participants in our study have Deaf families and consistent exposure to sign language, they participate in many activities bringing them into contact with hearing people, including relatives, teachers, neighbors, etc. Our expectation is that children will be closer to achieving discourse separation in the spoken language context, but not necessarily so in the sign context, as in the overall results of the study by Paradis and Nicoladis (2007).

Research Question 3: Do bimodal bilingual children match their language choice to that of their interlocutors? The null hypothesis is that there is no difference between children and their interlocutors. However, we expect children not to simply mirror their interlocutors, but to be influenced by a variety of variables in their language selection.

Research Question 4: Does the pattern of language selection change over time as children develop? We are particularly interested in the possibility that children increase their degree of language separation in the later stages of observation. However, it is possible that a fair amount of mixing will still be observed, since the oldest child in our study was still younger than the youngest child in the study by Paradis and Nicoladis (2007).

Research Question 5: Does the pattern of language selection vary for children in the U.S. compared with children in Brazil? Since our report involves four case studies, two from the U.S. and two from Brazil, we can begin to address the question of possible language-specific or culture-specific differences. However, it would be necessary to study a larger group of children to be able to definitively distinguish language or culture effects from individual differences.

METHODS

PARTICIPANTS

Participants are four male bimodal bilingual children and their adult interlocutors. The children are included in our long-term project, "Development of Bimodal Bilingualism," through which they have been involved in data collection with us over a period of years. For all of the children, the home language is a sign language (ASL/Libras); all four receive input in a spoken language (Eng/BP) through other relatives, neighbors, and the community.

- BEN (ASL/English) has two Deaf parents, one Deaf older sibling and one hearing older sibling, one Deaf grandparent and three hearing grandparents. His parents characterize the home environment as predominantly ASL, with some sign+speech blending.
- TOM (ASL/English) has two Deaf parents, and one hearing younger sibling, and no other Deaf family members with whom he has regular contact. His parents characterize the home environment as predominantly ASL, with some sign+speech blending.
- EDU (Libras/BP) has two Deaf parents. His mother signs and uses sign+speech blending. She understands his BP very well. His father only signs.

Table 1 | Participant information.

Child	Age range	Number speech sessions	Number sign sessions	Total number utterances produced			
				Child, speech sessions	Child, sign sessions	Interloc., speech sessions	Interloc., sign sessions
BEN (US)	1;04–3;06	8	8	4311 [3078]	2672 [1593]	4577 [3646]	3972 [2899]
TOM (US)	1;05–3;01	4	5	941 [432]	1200 [407]	1106 [712]	2121 [1261]
EDU (BR)	2;00–3;03	4	3	2652 [1866]	1180 [835]	2067 [1617]	899 [803]
IGOR (BR)	2;01–3;07	5	4	2764 [1899]	1741 [1100]	3142 [2918]	1657 [1202]

- IGOR (Libras/BP) has a Deaf father and a hearing mother who is fluent in Libras. They predominantly sign at home. The mother signs and blends sign+speech when the father is present, but when they are by themselves, she speaks with him. His father only signs.

For this article, we have analyzed a subset of the videos collected from each child, focusing on the age range (roughly) 1;06–3;06, as detailed in **Table 1**. The table provides the age, number of sessions, and number of utterances produced by the children and their interlocutors. In the table, two figures are given for number of utterances: the first figure includes all utterances; the second figure gives the number of utterances included in the analysis, excluding utterances consisting simply of interjections, uninterpretable speech/sign, single points, immediate imitations, etc.

DATA COLLECTION

Participants were video-taped to collect as natural a sample as possible of their ordinary language use. Generally, a target language is established for each session (Sign/Speech), and the target language alternated in weekly sampling sessions. Interlocutors for Sign-target sessions were generally the child's Deaf parent(s), with participation from Deaf (or in some cases, coda) research assistants interacting with the target child and/or behind the camera. Interlocutors for Speech-target sessions were generally a hearing research assistant (RA; all were known signers); in IGOR's case it was his mother. In some cases, a hearing signer research assistant was in the room during target Sign sessions, or a Deaf person (RA or parent) was in the room during Speech target sessions. This person generally stayed behind the camera and did not engage the child. More specifically, one hearing signer was in the room for two of BEN's early Sign target sessions, and two of TOM's early Sign target sessions, and his mother or a Deaf RA was in the room for five of BEN's Speech target sessions. We will explain how we took this into consideration in our analyses below. Our goal was to elicit natural language use and to observe any mixing that occurs; we did not try to enforce language separation (see Chen Pichler et al., 2013 for more detail about our filming methods).

DATA PROCESSING

Data processing took place in two steps. Our first step involved transcription of the speech and sign, to build up the corpora on which our analysis depends (Chen Pichler et al., 2010; Quadros et al., 2012, 2014). We subsequently added additional coding for specific research purposes.

Transcription

Transcription was done in our research laboratories following the procedures and conventions described in Chen Pichler et al. (2010). To summarize our procedure: The ELAN program (http://tla.mpi.nl/tools/tla-tools/elan/; Crasborn and Sloetjes, 2008) was used for all annotations. Our primary goal is to create an annotated video which can be searched and further annotated for particular research goals. First, hearing assistants transcribed the spoken language used by all participants in the video (the target Child, primary interlocutor Adult1, other Child*n* or Adult*n* interlocutors). Ordinary orthography was used with the addition of special symbols as needed. This initial transcription was checked by another research assistant, and any disagreement was resolved by discussion with at least one additional assistant when needed. Next, (near−)native sign assistants annotated the signing produced by all participants in the video. Glosses (Eng/BP) were used to annotate signs, supplemented by additional conventions shared by all transcribers. Both speech and sign annotations were checked again through additional steps of the process. Utterances were identified as speech and/or sign, with a relatively wide net including all potential linguistic expressions in the initial transcription. Utterance breaks were determined using prosodic information as well as propositional information. Finally, a Free Translation tier was constructed taking into account both the sign language and the spoken language.

Coding

For the present analysis, coding required adding the following tiers to our basic ELAN template, with a set for each participant (Child, Adult1, and additional Adult*n* as needed):

- Modality: modality of the utterance (Sign (only), Speech (only), Bimodal, or Excluded)
- Interlocutor: the addressee (Deaf adult, hearing adult, parent, or target child)

Utterances were Excluded if they were completely unintelligible, or consisted of only spoken or signed routines, interjections, non-speech communicative vocalizations or non-sign communicative actions (gestures), or complete imitations of the interlocutor's immediately previous utterance—with no other speech or sign. For example, a spoken "well," "yes," "no," a head nod, an "oops" gesture, or a clap, if occurring by itself, was Excluded. For the Modality analysis, utterances were also Excluded if modality could

not be determined; e.g., there was audible speech but the speaker's hands were off-camera.

To count as "bimodal," we required that some portion of an utterance be presented in sign, and some portion in the spoken language, whether full voice or whispered. It was possible for us to clearly distinguish between full voice and whispering vs. mouthing in the auditory component of our recordings. In whispering, there is turbulence during speech which is not present during mouthing. We did not count mouthing with sign as bimodal (unlike van den Bogaerde, 2000; van den Bogaerde and Baker, 2005, 2009; Baker and van den Bogaerde, 2008). Mouthing with sign in ASL and Libras is quite variable. Mouthing is considered by some to be a mark of influence from the spoken language, and an indication that "contact signing" is being used (cf. Lucas and Valli, 1992). However, many Deaf native signers use mouthing frequently, and more and more linguistic analyses have treated mouthing as a part of the sign language (e.g., Nadolske and Rosenstock, 2007). From our perspective, mouthing may sometimes be a mark of bilingualism, but it is also sometimes a part of signing. Our decision to require full voice or whispering for an utterance to qualify as bimodal obviates the need to judge the status of specific instances of voiceless mouthing. Of course, it means that our figures for proportion of blending cannot be directly compared with those of researchers who include mouthing (e.g., van den Bogaerde and Baker references cited above).

In the initial analysis, combinations of speech and sign interjections (e.g., spoken "yes" with a head nod) were counted as Bimodal, as were combinations of speech with only an index/pointing sign. In subsequent analyses, such combinations were not included in bimodal counts.

For each included utterance, interlocutor was determined by examination of the video. In most cases, the child is addressing the primary interlocutor and vice-versa. Occasionally, a different interlocutor is addressed; for example, the interlocutor might address the cameraperson to check the status of the cordless microphone. In some cases, more than one interlocutor is present, such as when the child is filmed with both Deaf parents.

All of the Brazilian data was coded once by a single coder, and checked and modified by a second coder. Most of the US data

was coded by a single coder, with another experienced coder providing coding for a small subset of the data. To check reliability, 5% of the US data were coded blind by a second coder. After the two codings were compared by a third experienced coder, it was determined that accuracy of coding modality was over 93%, and interlocutor coding was over 97% accurate.

RESULTS

OVERALL ANALYSIS

For our first analysis, we calculated the proportion of sign, speech, and bimodal utterances produced by the children and their interlocutors across all contexts within speech target sessions and sign target sessions. The results of this calculation are presented in **Table 2**. Two things are immediately clear. First, the children showed differentiated production in Speech vs. Sign target sessions. This is confirmed by a series of four 2×3 chi-square tests of independence (n ranged from 943 to 4671, $\chi^2 = 163.5\text{--}1512.58$, $p < 0.0001$ for all four tests, Cramer's $V = 0.3123\text{--}0.5683$). Second, the children were distinct from their interlocutors in their patterns of speech, bimodal, and sign production in both Speech and Sign target sessions (for seven of eight chi-square tests, $n = 1356\text{--}6724$, $\chi^2 = 54.21\text{--}1130.18$, $p < 0.0001$, Cramer's $V = 0.1574\text{--}0.8128$; the effect is marginal at $\chi^2_{(2, n = 4813)} = 5.82$, $p = 0.0545$, Cramer's $V = 0.0348$ for the comparison between IGOR's output and that of his interlocutors in Speech sessions).

DEVELOPMENTAL ANALYSIS

To further investigate the patterns of language mixing by bimodal bilinguals, as distinct from other hearing children, we conducted a second analysis in which we eliminated bimodal utterances where speech was accompanied by pointing, but no other sign. Such speech+pointing combinations are not unique to bimodal bilingual children, as they are commonly reported in studies of hearing, non-signing children (Capirci et al., 1996; Ozçaliskan and Goldin-Meadow, 2005), where points accompanying speech are classified as gesture. Our elimination of speech+point combinations from the second analysis was a conservative measure, given the considerable debate in the sign linguistics field over the status of pointing in sign language. For the same reason, we also excluded combinations consisting solely of elements that would be Excluded if occurring alone (e.g., sign+speech interjections or speech+gesture).

Table 2 | Overall results.

		BEN		TOM		EDU		IGOR	
		Child	**Interlocutor**	**Child**	**Interlocutor**	**Child**	**Interlocutor**	**Child**	**Interlocutor**
Speech target	Speech	0.58	0.73	0.81	0.75	0.93	0.74	0.78	0.75
	Bimod.	0.37	0.24	0.13	0.24	0.06	0.21	0.18	0.20
	Sign	0.05	0.03	0.06	0.00	0.01	0.05	0.04	0.05
Sign target	Speech	0.11	0.05	0.41	0.01	0.65	0.02	0.51	0.03
	Bimod.	0.43	0.10	0.36	0.10	0.17	0.03	0.29	0.08
	Sign	0.45	0.85	0.23	0.89	0.18	0.96	0.20	0.89

Proportion of speech, bimodal and sign utterances produced by each child and interlocutors in Speech target sessions and Sign target sessions.
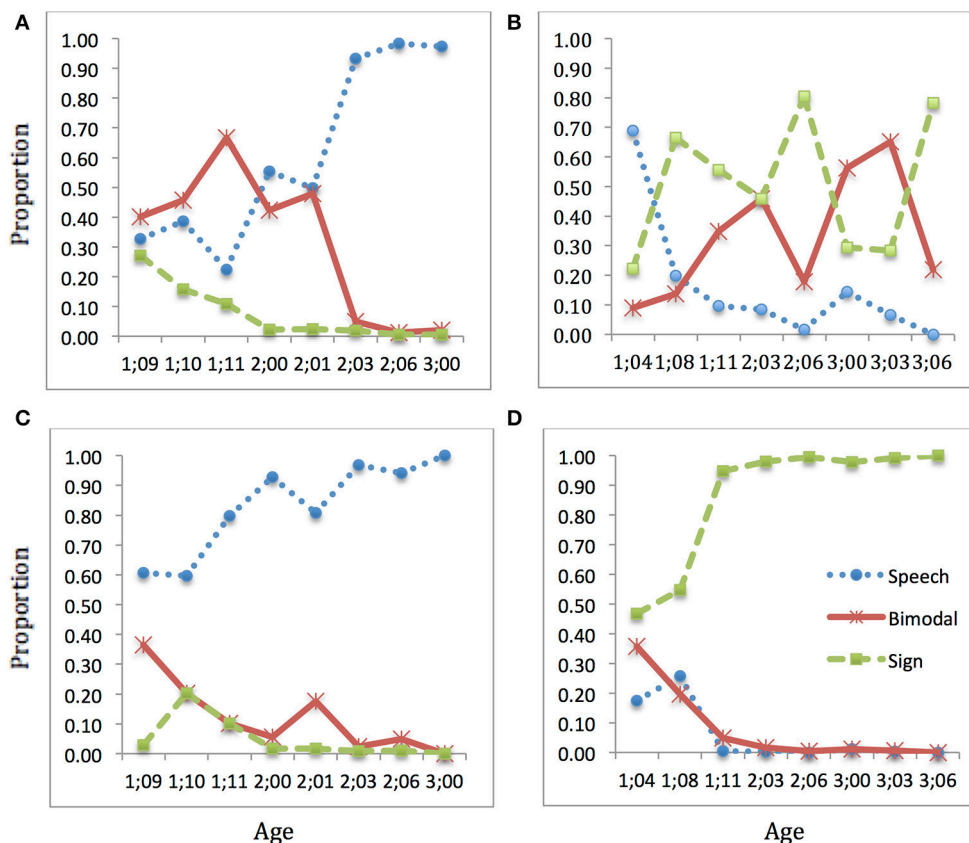
FIGURE 5 | Language choice over time: BEN. Proportion of speech (blue dotted), sign (green dashed), and bimodal (red solid) utterances produced: **(A)** child speech target; **(B)** child sign target; **(C)** interlocutor(s) speech target; **(D)** interlocutor(s) sign target (Note: Intervals along x-axis are not regular).

In addition, we separated out utterances addressed to different interlocutors. In particular, the US sessions occasionally included multiple interlocutors with different auditory status. We focused on the children's productions to hearing interlocutors in the Speech sessions and to Deaf interlocutors in the Sign sessions. We also focused on the interlocutors' utterances to the target child, excluding those addressed to other participants. Finally, we calculated the proportion of speech, bimodal, and sign productions at each session, in order to observe possible developmental effects. The results of these calculations are displayed graphically in **Figures 5–8**.

A series of chi-square tests were applied to see whether the modality pattern (speech, bimodal, sign) produced by each child was different from that produced by the interlocutor(s) in the same session. A second series of chi-square tests examined the difference between each child's own productions in Speech target sessions and Sign target sessions at comparable ages. A full table of the results of these comparisons is available in the Supplementary Materials for this article. The results are summarized in **Table 3**.

To summarize: with very few exceptions, virtually every comparison showed a significant difference between each child and his interlocutors, and between each child's own productions in speech and sign target sessions.

DISCUSSION

Let us interpret the results of our analyses within the context of the five research questions raised in the Section called Expectations for the Current Study.

Research Question 1: Do developing bimodal bilingual children show interlocutor sensitivity by selecting language modality at differential rates in Speech- and Sign-target sessions? In particular, do they produce a greater proportion of spoken language in Speech-target sessions and a greater proportion of sign language in Sign-target sessions? Our prediction was that, counter to the null hypothesis, children would differ in language selection across different target language sessions.

Our expectation in this case was confirmed by our overall analysis presented in **Table 2**. The children did indeed differ in their language selection across contexts, with each child producing more speech in the Speech-target sessions than in the Sign-target sessions, and more sign in the Sign-target sessions than in the Speech-target sessions.

Research Question 2: Is any difference between Speech-target sessions and Sign-target sessions influenced by the dominance of the spoken language in the broader sociolinguistic context? Our expectation was that children would be closer to achieving discourse separation in the spoken language context, but not necessarily so in the sign context. This expectation was also
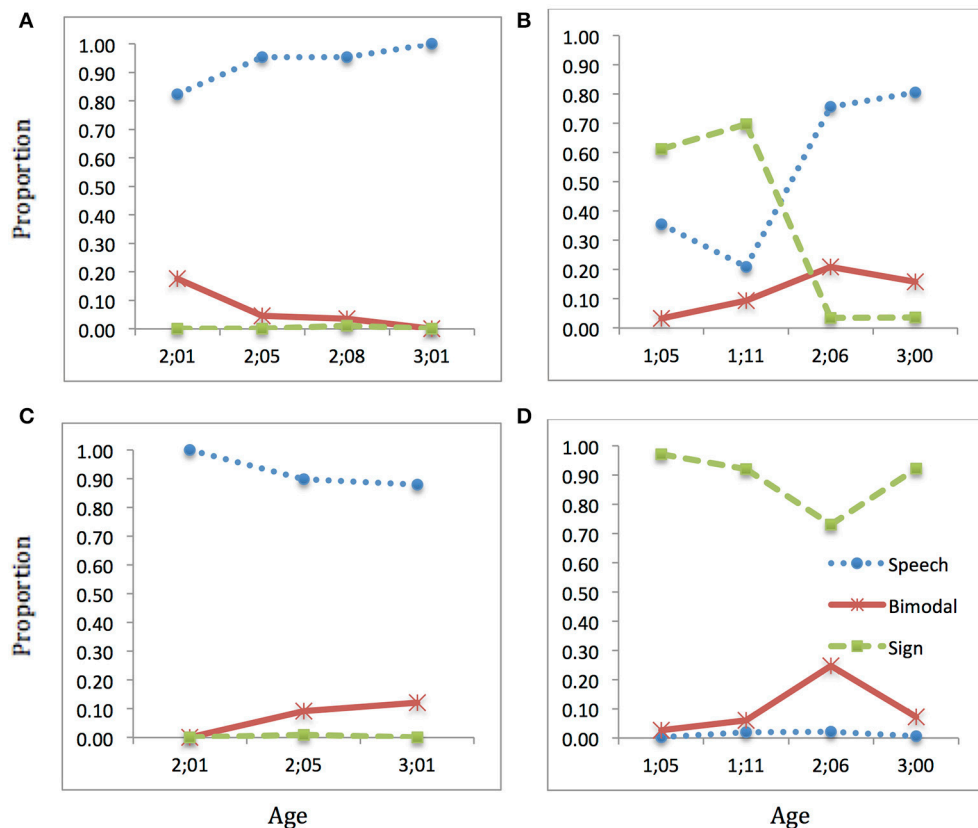
FIGURE 6 | Language choice over time: TOM. Proportion of speech (blue dotted), sign (green dashed), and bimodal (red solid) utterances produced: **(A)** child speech target; **(B)** child sign target; **(C)** interlocutor(s) speech target; **(D)** interlocutor(s) sign target (Note: Intervals along x-axis are not regular).

borne out. Looking at the results in **Table 2** again, we see that each child had a higher proportion of speech in the Speech-target sessions than their proportion of sign in the Sign-target sessions. In the overall analysis, three of the children (TOM, EDU, and IGOR) had over 75% use of speech in Speech sessions, but less than 25% use of sign in Sign target sessions. In this respect, their overall performance was comparable to the degree of language separation exhibited by three of the English-dominant children in the study by Paradis and Nicoladis (2007; cf. Figure 2).

There is a possible alternative explanation for our observation that children were closer to achieving discourse separation in the spoken language context than in the sign context. Rather than a function of the strong dominance of spoken language in the broader sociolinguistic context, this finding could be due to some kind of special tuning of the human linguistic system that preferences speech over sign. The possibility that the human linguistic system preferences speech might be supported by the observation that sign languages are reserved for contexts in which spoken language won't do—Deaf communities, and hearing communities that for various reasons don't speak (e.g., certain religious orders, or persons working in very loud conditions). To the contrary, some researchers have explicitly argued that the human linguistic system is amodal, equipotential for input in a sign language or a spoken language (Petitto and Marentette, 1991).

If the human linguistic system has a preference for spoken language, we might well expect hearing people to uniformly show this preference, despite having input in a sign language from birth. They might even be expected to have difficulty switching from the preferred, dominant language to the less preferred one.

Indeed, Emmorey et al. (2008b; 2013) find that their Coda participants in general show dominance in speech, based on self-report of skill level, and psycholinguistic task responses. However, these findings represent the participants in their experiments as a whole. It is not the case that every individual showed the same pattern, and the self-report ratings for proficiency in speech vs. sign are very close. One participant in the study by Emmorey et al. (2008a) responded to one of the tasks using ASL only. In addition, anecdotal reports by adult codas indicate that many consider ASL to be their primary language (Bishop, 2010).

Even if signed and spoken language are equipotential (Petitto, 1994), it would not be surprising to find a strong tendency for hearing native signers to be (or become) speech dominant. Even for those who work in an environment with others using sign language (e.g., a school), a truly balanced or sign-dominant environment would be rare. In the absence of a method to control for or counterbalance such a potentially overwhelming factor, data on the dominance of speech vs. sign in bimodal bilinguals will not be able to rule out (or support) the hypothesis that an overall linguistic preference for speech is at work. Nevertheless, taking into
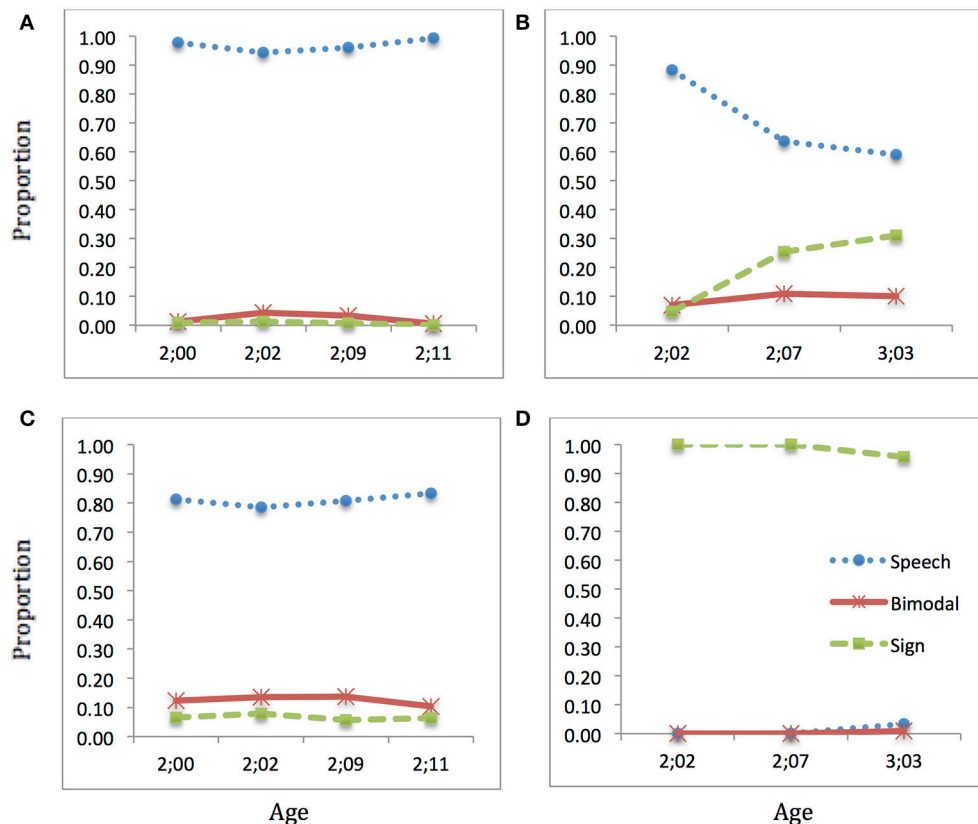
FIGURE 7 | Language choice over time: EDU. Proportion of speech (blue dotted), sign (green dashed), and bimodal (red solid) utterances produced: **(A)** child speech target; **(B)** child sign target; **(C)** interlocutor(s) speech target; **(D)** interlocutor(s) sign target (Note: Intervals along x-axis are not regular).

consideration individual differences in the strength of the asymmetry between speech and sign, we will continue to consider the environment as a primary causal factor.

Research Question 3: Do bimodal bilingual children match their language choices to that of their interlocutors? The null hypothesis is that there is no difference between children and their interlocutors, but our overall analysis revealed a significant difference (**Table 2**). Of the eight comparisons between the four children and their interlocutors in Speech and Sign sessions, seven were highly significant and one (IGOR speech) was marginal. However, visual inspection of the numbers in **Table 2** makes it clear that the patterns of usage for the children are much closer to those of their interlocutors in the Speech sessions than in the Sign sessions. In addition, the values for Cramer's V are much higher for the Sign sessions (range 0.4222–0.8128) than for the Speech sessions (range 0.1574–0.2623 for the three significant results), indicating that the differences between the children and their interlocutors are much higher in the Sign sessions.

We take these results as a strong indication that children's language choice is a function of their developing knowledge of the two languages and their appropriate contexts of usage. We will return to this point in the discussion of Research Question 4.

Research Question 4: Does the pattern of language selection change over time as children develop? For this question, we refer to the developmental analysis presented in the Section called Developmental Analysis. As the graphs indicate, the children's choices did change over time, but in different ways for each individual child. We discuss each child's results in turn.

BEN

The results presented in **Figure 5** show that in Speech sessions, BEN's use of sign started relatively low and declined quickly to essentially zero by age 2;00, but his use of bimodal utterances continued along with speech until 2;03, from which point he achieved complete discourse separation for Speech. It is interesting to note that his interlocutors' use of sign and bimodal utterances was also relatively high during the earliest sessions, with sign reaching zero by 2;00 and bimodal by 2;03. In this respect, BEN and his interlocutors were similar, but it is not clear whether it was BEN's use of sign and bimodal productions that encouraged the interlocutors to use these modalities or vice-versa. We note that at 2;03, the statistical comparison did not show a significant difference between BEN and his Speech interlocutors (it was marginal at $p = 0.063$), and at the two later ages, the chi-square test could not be done because of low expected frequencies in two cells—this in turn being due to the low use of sign by both BEN and interlocutors. Thus, he clearly moved toward the same pattern of production in Speech sessions (speech only) as his interlocutors did.
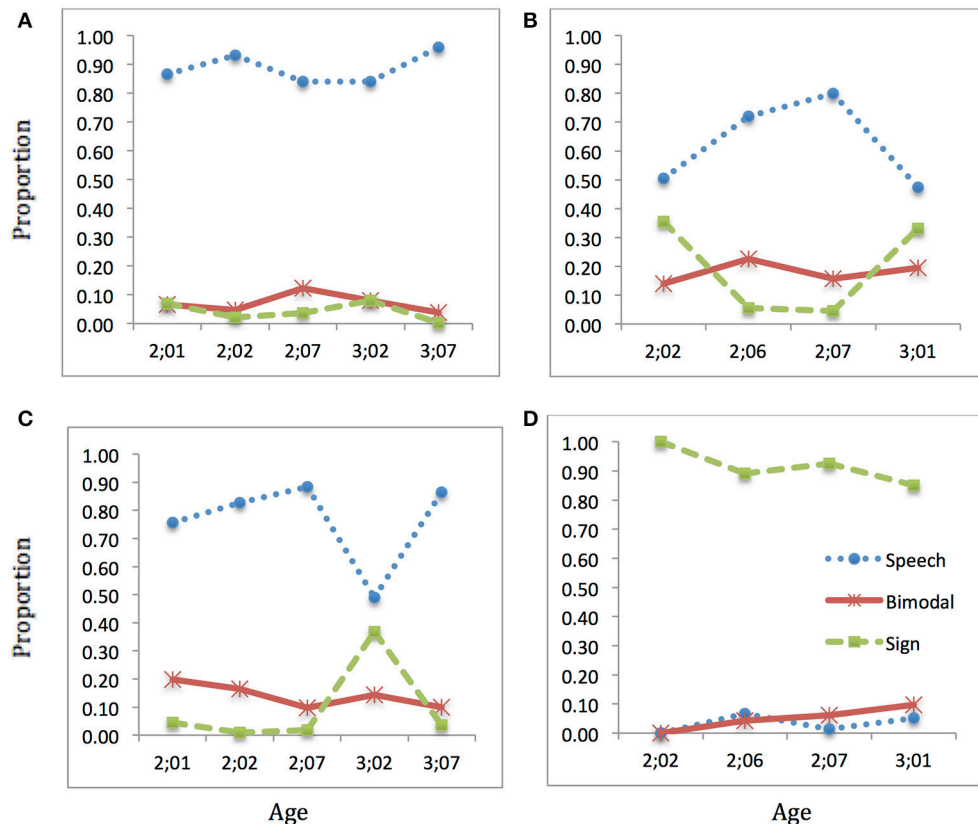
FIGURE 8 | Language choice over time: IGOR. Proportion of speech (blue dotted), sign (green dashed), and bimodal (red solid) utterances produced: **(A)** child speech target; **(B)** child sign target; **(C)** interlocutor(s) speech target; **(D)** interlocutor(s) sign target (Note: Intervals along x-axis are not regular).

The picture is quite different, and very interesting, in BEN's Sign sessions. First, we observe that BEN's mother (the primary interlocutor in all but the last one of the Sign sessions reported here) made a notable change in her own productions. In fact, she reported to us that she originally thought it would be best to use blending with her hearing child, but she decided when he was 1;11 to stop using speech with him and use sign exclusively. The data from our observation sessions indicate that she adhered to this commitment. BEN's own use of speech in Sign sessions decreased dramatically over the early period, and reached a low baseline by 1;11. However, BEN did not use sign exclusively while not using speech in Sign sessions; rather, he used a mixture of sign and bimodal productions. The proportion of sign and bimodal production fluctuated greatly, with no apparent pattern.

As mentioned earlier, in the US sessions occasionally additional research assistants were present in the room but not interacting directly with the children. In order to see whether the presence of other adults in the room affected children's use of speech vs. sign, we checked carefully to see which sessions had additional participants (e.g., a camera-person) and how this relates to language choice. For BEN's Sign sessions, a hearing (signer) was present in the first three sessions only. All of the later sessions—in particular, those showing great fluctuations in the use of sign vs. bimodal productions—had only Deaf people

present in the room. There was greater variability in the Speech sessions, with a Deaf person present in five of the eight sessions throughout the observation period. However, as noted, BEN became quite dominant in using speech only during Speech sessions, apparently despite the occasional presence of a Deaf person in the room.

TOM

TOM's pattern of results, presented in **Figure 6**, show that he had a high and increasing tendency to use speech only in Speech sessions throughout the observation period. His hearing interlocutors showed a slight trend in the opposite direction, using more bimodal productions over time. It is this difference that likely led to the overall significant difference between TOM and his interlocutors in Speech sessions, even though the differences were not significant in the two earliest sessions (note that at 2;01, neither TOM nor his interlocutor produced very many utterances that could be included in the second analysis, because there were multiple participants in the session and we had to exclude many utterances). There were only hearing interlocutors and people in the room during the Speech sessions coded.

In Sign sessions, a hearing signer was present at the earliest session, but otherwise only Deaf people were present. TOM's interlocutors predominantly used sign, but there was an increase

Table 3 | Developmental results.

		Number comparisons	Number comparisons significant at $p < 0.05$	Number comparisons significant at $p < 0.0001$
BEN	Speech vs. input[a]	6	5	5
	Sign vs. input[b]	7	7	6
	Speech vs. sign	4	4	4
TOM	Speech vs. input[b]	2	1	1
	Sign vs. input[a]	2	2	2
	Speech vs. sign[a]	1	1	1
EDU	Speech vs. input	4	4	4
	Sign vs. input	3	3	3
	Speech vs. sign	2	2	2
IGOR	Speech vs. input	5	4	4
	Sign vs. input	4	4	4
	Speech vs. sign	3	3	2

Comparisons between Child vs. Input for Speech, Child vs. Input for Sign, Child Speech vs. Child Sign.

[a] *Two additional comparisons could not be conducted because two or more expected cell frequencies were calculated to be smaller than 5.*

[b] *One additional comparison could not be conducted because two or more expected cell frequencies were calculated to be smaller than 5.*

in bimodal productions by his mother, the primary interlocutor in the session at age 2;06. Our informal observations suggest that TOM's mother did use bimodal productions with him and with other people often, so we do not take this to be a misrepresentation of his input in general. TOM's own productions in Sign sessions displayed an increase over time in speech and bimodal production, with a corresponding decrease in sign. Thus, by 3;00 TOM showed a strong tendency to use speech in both types of sessions, while still distinguishing between the two contexts.

EDU

EDU's pattern of language selection, shown in **Figure 7**, showed little change over time in Speech sessions. His use of speech was at ceiling in these sessions, despite the notably lower rate of speech and correspondingly higher rate of bimodal and sign productions by his interlocutors. In Sign sessions, EDU started with a high proportion use of speech, but this was moderated over time, moving toward higher use of Sign but relatively low use of bimodal productions. The interlocutors in his Sign sessions—his Deaf mother and father—used sign almost exclusively on camera, but we observed that his mother used speech/bimodal productions with him and with others at other times. Overall, EDU showed a strong speech bias in the observations presented here.

IGOR

IGOR's developmental data, shown in **Figure 8**, revealed a fairly constant, high use of speech in Speech sessions. Like EDU, IGOR used more speech than his interlocutors, who also made use

of bimodal productions (with an inexplicable increase in the number of sign productions in one session, at 3;02).

In Sign sessions, IGOR used a mixture of speech, sign, and bimodal productions. He appeared to be increasing the amount of sign and correspondingly decreasing the amount of speech by the end of the observation period (3;01). His interlocutors used a high proportion of sign productions, with some bimodal production as well.

Although the details of his production were different, IGOR appears overall to be similar to EDU in showing a strong preference for speech, with movement toward more use of sign and bimodal production in Sign sessions after age 3.

Research Question 5: Does the pattern of language selection vary for children in the U.S. compared with children in Brazil? Since our report involves only four case studies, it is difficult to definitively distinguish language or culture effects from individual differences. Overall, our impression was that TOM, EDU and IGOR showed similarities in performance, as children who favor spoken language and therefore display discourse separation most clearly for their spoken language, but also interlocutor sensitivity for their sign language. Only BEN showed evidence of complete discourse separation for both languages, but this is likely to be an individual difference. No clear language/culture effects were thus observed in our data.

CONCLUSIONS

According to the model of bimodal bilingualism we presented in **Figure 1**, bilinguals have the option of using grammatical knowledge and lexical items from either language, separately or in combination, as long as general constraints on language structure are met. Further constraints on the use of code-mixing (including code-blending) may be imposed by the sociolinguistic environment: some communities take more advantage of the mixing available to bilinguals, while others tend to avoid it. Thus, children must learn to take into consideration both the structural properties afforded by their languages and the language usage patterns exhibited by individual interlocutors and language communities.

The children in our study showed that they are sensitive to the language used by interlocutors, in that they displayed differential language selection in Speech– vs. Sign-target sessions. Three of the four participants were also strongly affected by the dominance of the spoken language in the broader sociolinguistic community: they distinguished between Speech and Sign contexts, yet showed a preference for use of speech in both contexts. The fourth participant, BEN, showed a full discourse separation pattern, if we count his use of bimodal productions as "appropriate" for the Sign sessions.

One might ask why BEN would use bimodal productions rather than exclusively using sign in Sign-target sessions, given his apparent facility and recognition of the role of sign. Emmorey et al. (2008a) and Pyers and Emmorey (2008), observing adult codas, proposed that codas use code-blending, and even use aspects of ASL non-manual marking while speaking English to non-signers, because (complete) inhibition or suppression of the unselected language has a processing cost.

For unimodal bilinguals, use of one language necessitates inhibition of the other language, whereas bimodal bilinguals can use blending to ease the burden of inhibition to varying extents. We suggest that this tendency to blend when inhibition is difficult lies behind BEN's use of blending in the Sign-target sessions. The same might be true for the other three participants, but their rates of blending were overall lower than BEN's.

While the children all showed interlocutor sensitivity, they did not mirror their interlocutors' rates of production of speech, sign, and bimodal utterances. Still, it is quite possible that the attitude of the children's input providers played a role in their language selection, as suggested by Döpke (1992) and Lanza (1997) for unimodal bilinguals, and by van den Bogaerde and Baker (2009) for NGT-Dutch bimodal bilinguals and Kanto et al. (2013) for FinSL-Finnish bimodal bilinguals. In general, the children in our study are exposed to blending from at least one parent, with relatively less sign-only input, and all of the Deaf parents are bilingual to some degree, whether or not they use speech with their hearing child. Many of them also show their understanding of their children's spoken output: for example, EDU's mother answers (in sign) his spoken questions, showing that he can achieve successful communication with her even when he uses speech. In addition, during our data collection sessions the children interacted with numerous hearing people who are known signers, and they also modeled the use of blending. The only case we know where a stricter monolingual strategy is pursued is BEN's mother.

Additional research would be needed to confirm this, but our overall findings are in agreement with those researchers suggesting that greater discourse separation is related to greater adherence to a monolingual strategy. In addition, as Chen Pichler et al. (2014) discuss, maintenance of a minority home language for kodas may be supported through increased opportunities for them to use that language with a variety of interlocutors, including peers, throughout development.

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://www.frontiersin.org/journal/10.3389/fpsyg.2014.01163/abstract

REFERENCES

Baker, A., and van den Bogaerde, B. (2008). "Code-mixing in signs and words in input to and output from children," in *Sign Bilingualism: Language Development, Interaction, and Maintenance in Sign Language Contact Situations*, eds C. Plaza Pust and E. Morales Lopez (Amsterdam: John Benjamins), 1–28.

Baker, A., and van den Bogaerde, B. (2014). "KODAs: a special form of bilingualism," in *Multilingual Aspects of Signed Language Communication and Disorder*, ed D. Quinto-Pozos (Bristol: Multilingual Matters), 211–234.

Bhatia, T. K., and Ritchie, W. C. (1999). "The bilingual child: some issues and perspectives," in *Handbook of Child Language Acquisition*, eds W. C. Ritchie and T. K. Bhatia (San Diego, CA: Academic Press), 569–643.

Bhatt, R., and Bolonyai, A. (2011). Code-switching and the optimal grammar of bilingual language use. *Biling. Lang. Cogn.* 14, 522–546. doi: 10.1017/S1366728910000295

Bishop, M. (2010). Happen can't hear: an analysis of code-blends in hearing, native signers of American Sign Language. *Sign Lang. Stud.* 11, 205–240. doi: 10.1353/sls.2010.0007

Bishop, M., and Hicks, S. (2005). Orange eyes: bimodal bilingualism in hearing adults from Deaf families. *Sign Lang. Stud.* 5, 188–230. doi: 10.1353/sls.2005.0001

Capirci, O., Iverson, J. M., Pizzuto, E., and Volterra, V. (1996). Communicative gestures during the transition to two-word speech. *J. Child Lang.* 23, 645–673. doi: 10.1017/S0305000900008989

Chen Pichler, D., Hochgesang, J., Lillo-Martin, D., and Quadros, R. M. (2010). Conventions for sign and speech transcription of child bimodal bilingual corpora in ELAN. *LIA* 1, 11–40. doi: 10.1075/lia.1.1.03che

Chen Pichler, D., Hochgesang, J., Lillo-Martin, D., Quadros, R. M., and Reynolds, W. (2013). *Best Practices for Building a Bimodal Bilingual Child Language Corpus*. Avaliable online at: http://bibibi.uconn.edu/Publications.html

Chen Pichler, D., Lee, J., and Lillo-Martin, D. (2014). "Language development in ASL-english bimodal bilinguals," in *Multilingual Aspects of Signed Language Communication and Disorder*, ed David Quinto-Pozos (Bristol: Multilingual Matters), 235–260.

Chomsky, N. (1995). *The Minimalist Program*. Cambridge: The MIT Press.

Comeau, L., Genesee, F., and Lapaquette, L. (2003). The modeling hypothesis and child bilingual codemixing. *Int. J. Biling.* 7, 113–126. doi: 10.1177/13670069030070020101

Costa, A., and Santesteban, M. (2004). Lexical access in bilingual speech production: evidence from language switching in highly proficient bilinguals and L2 learners. *J. Mem. Lang.* 50, 491–511. doi: 10.1016/j.jml.2004.02.002

Crasborn, O., and Sloetjes, H. (2008). "Enhanced ELAN functionality for sign language corpora," in *Proceedings of LREC 2008, Sixth International Conference on Language Resources and Evaluation*, 39–43. Available online at: http://www.lrec-conf.org/proceedings/lrec2008

Deuchar, M., and Quay, S. (1999). Language choice in the earliest utterances: a case study with methodological implications. *J. Child Lang.* 26, 461–475. doi: 10.1017/S0305000999003852

Donati, C., and Branchini, C. (2013). "Challenging linearization: simultaneous mixing in the production of bimodal bilinguals," in *Challenges to Linearization*, eds T. Biberauer and I. Roberts (Berlin: Mouton De Gruyter), 93–128.

Döpke, S. (1992). One *Parent—One Language: An Interactional Approach*. Amsterdam: John Benjamins.

Emmorey, K., Borinstein, H. B., Thompson, R., and Gollan, T. H. (2008a). Bimodal bilingualism. *Biling. Lang. Cogn.* 11, 43–61. doi: 10.1017/S1366728907003203

Emmorey, K., Luk, G., Pyers, J. E., and Bialystok, E. (2008b). The source of enhanced cognitive control in bilinguals: evidence from bimodal bilinguals. *Psychol. Sci.* 19, 1201–1206. doi: 10.1111/j.1467-9280.2008.02224.x

Emmorey, K., Petrich, J. A. F., and Gollan, T. H. (2013). Bimodal bilingualism and the frequency-lag hypothesis. *J. Deaf Stud. Deaf Educ.* 18, 1–11. doi: 10.1093/deafed/ens034

Fung, C. (2012). *Code-Blending in Early Hong Kong Sign Language: A Case Study*. M.A. thesis, The Chinese University of Hong Kong.

Genesee, F. (1989). Early bilingual development: one language or two? *J. Child Lang.* 16, 161–179. doi: 10.1017/S0305000900013490

Genesee, F., Nicoladis, E., and Paradis, J. (1995). Language differentiation in early bilingual development. *J. Child Lang.* 22, 611–631. doi: 10.1017/S0305000900009971

Gonzalez-Vilabazo, K., and López, L. (2012). Little v and parametric variation. *Nat. Lang. Linguist. Theory* 30, 33–77. doi: 10.1007/s11049-011-9141-5

Griffith, P. L. (1985). Mode-switching and mode-finding in a hearing child of deaf parents. *Sign Lang. Stud.* 48, 195–222. doi: 10.1353/sls.1985.0010

Grosjean, F. (1989). Neurolinguist, beware! The bilingual is not two monolinguals in one person. *Brain Lang.* 36, 3–15. doi: 10.1016/0093-934X(89)90048-5

Halle, M., and Marantz, A. (1993). "Distributed morphology and the pieces of inflection," in *The View from Building 20: Essays in Honor of Sylvain Bromberger*, eds K. Hale and S. J. Keyser (Cambridge, MA: MIT Press), 111–176.

Kanto, L., Huttunen, K., and Laasko, M.-L. (2013). Relationship between the linguistic environments and early bilingual language development of hearing children in Deaf-parented families. *J. Deaf Stud. Deaf Educ.* 18, 242–260. doi: 10.1093/deafed/ens071

Koulidobrova, E. (2012). *When the Quiet Surfaces: 'Transfer' of Argument Omission in the Speech of ASL-English Bilinguals*. Ph.D. dissertation, University of Connecticut.

Kroll, J. F., Bobb, S. C., and Wodniecka, Z. (2006). Language selectivity is the exception, not the rule: arguments against a fixed locus of language selection in bilingual speech. *Biling. Lang. Cogn.* 9, 119–135. doi: 10.1017/S1366728906002483

Kroll, J. F., and Stewart, E. (1994). Category interference in translation and picture naming: evidence for asymmetric connections between bilingual memory representations. *J. Mem. Lang.* 33, 149–174. doi: 10.1006/jmla.1994.1008

Lanza, E. (1997). *Language Mixing in Infant Bilingualism*. New York, NY: Oxford University Press.

Lillo-Martin, D., Koulidobrova, H., Quadros, R. M., and Chen Pichler, D. (2012). "Bilingual language synthesis: evidence from WH-questions in bimodal bilinguals," in *Proceedings of the 36th Annual Boston University Conference on Language Development*, eds A. K. Biller, E. Y. Chung, and A. E. Kimball (Somerville, MA: Cascadilla Press), 302–314.

Lillo-Martin, D., Quadros, R. M., Koulidobrova, H., and Chen Pichler, D. (2010). "Bimodal bilingual cross-language influence in unexpected domains," in *Language Acquisition and Development: Proceedings of GALA 2009*, eds J. Costa, A. Castro, M. Lobo, and F. Pratas (Newcastle upon Tyne: Cambridge Scholars Press), 264–275.

Lucas, C., and Valli, C. (1992). *Language Contact in the American Deaf Community*. San Diego, CA: Academic Press.

MacSwan, J. (2000). The architecture of the bilingual language faculty: evidence from intrasentential code switching. *Biling. Lang. Cogn.* 3, 37–54. doi: 10.1017/S1366728900000122

MacSwan, J. (2005). Codeswitching and generative grammar. *Biling. Lang. Cogn.* 8, 1–22. doi: 10.1017/S1366728904002068

Nadolske, M. A., and Rosenstock, R. (2007). "Occurrence of mouthings in American Sign Language: a preliminary study," in *Visible Variation: Comparative Studies on Sign Language Structure*, eds P. M. Perniss, R. Pfau, and M. Steinbach (Berlin: Walter de Gruyter), 35–62.

Nicoladis, E., and Genesee, F. (1996). A longitudinal study of language differentiation in young bilingual children. *Lang. Learn.* 46, 439–464. doi: 10.1111/j.1467-1770.1996.tb01243.x

Nicoladis, E., and Secco, G. (2000). The role of a child's productive vocabulary in the language choice of a bilingual family. *First Lang.* 20, 003–028. doi: 10.1177/014272370002005801

Ozçaliskan, S., and Goldin-Meadow, S. (2005). Gesture is at the cutting edge of early language development. *Cognition* 96, B101–B113. doi: 10.1016/j.cognition.2005.01.001

Paradis, J. (2001). Do bilingual two-year-olds have separate phonological systems? *Int. J. Biling.* 5, 19–38. doi: 10.1177/13670069010050010201

Paradis, J., and Nicoladis, E. (2007). The influence of dominance and sociolinguistic context on bilingual preschoolers' language choice. *Int. J. Biling. Educ. Biling.* 10, 277–297. doi: 10.2167/beb444.0

Petitto, L. A. (1994). "On the equipotentiality of signed and spoken language in early language ontogeny," in *Post-Milan ASL and English Literacy: Issues, Trends,*

and Research, ed B. Snider (Washington, DC: Gallaudet University Press), 195–223.

Petitto, L. A., Katerelos, M., Levi, B., Gauna, K., Tetrault, K., and Ferraro, V. (2001). Bilingual signed and spoken language acquisition from birth: implications for the mechanisms underlying early bilingual language acquisition. *J. Child Lang.* 28, 453–496. doi: 10.1017/S0305000901004718

Petitto, L. A., and Marentette, P. (1991). Babbling in the manual mode: evidence for the ontogeny of language. *Science* 251, 1483–1496. doi: 10.1126/science.2006424

Petroj, V., Guerrera, K., and Davidson, K. (2014). "ASL dominant code-blending in the whispering of bimodal bilingual children," in *Proceedings of the 36th Annual Boston University Conference on Language Development* (Sommerville, MA: Cascadilla Press).

Poplack, S. (1980). Sometimes I'll start a sentence in Spanish y termino en español: toward a typology of code-switching. *Linguistics* 18, 581–618. doi: 10.1515/ling.1980.18.7-8.581

Pyers, J., and Emmorey, K. (2008). The face of bimodal bilingualism: grammatical markers in American Sign Language are produced when bilinguals speak to English monolinguals. *Psychol. Sci.* 19, 531–535. doi: 10.1111/j.1467-9280.2008.02119.x

Quadros, R. M., Lillo-Martin, D., and Chen Pichler, D. (2012). "Bi-National Bi-modal Bi-Lingual corpora of child language," in *Proceedings of the VIIth GSCP (Gruppo di Studi sulla Comunicazione Parlata) International Conference: Speech and Corpora*, eds H. Mello, M. Pettorino, and T. Raso (Firenze: Firenze University Press), 87–92.

Quadros, R. M., Lillo-Martin, D., and Chen Pichler, D. (2013). "Early effects of bilingualism on WH-question structures: insight from sign-speech bilingualism," in *Proceedings of GALA 2011*, eds S. Stavrakaki, M. Lalioti, and P. Konstantinopoulou (Newcastle upon Tyne: Cambridge Scholars Press), 300–308.

Quadros, R. M., Lillo-Martin, D., and Chen Pichler, D. (2014). "Methodological considerations for the development and use of sign language acquisition data," in *Spoken Corpora and Linguistic Studies*, eds T. Raso, H. Mello, and M. Pettorino (Amsterdam: John Benjamins Publishers), 84–102.

Shook, A., and Marian, V. (2012). Bimodal bilinguals co-activate both languages during spoken comprehension. *Cognition* 124, 314–324. doi: 10.1016/j.cognition.2012.05.014

Unsworth, S. (2013). Current issues in multilingual first language acquisition. *Annu. Rev. Appl. Linguist.* 33, 21–50. doi: 10.1017/S0267190513000044

van den Bogaerde, B. (2000). *Input and Interaction in Deaf Families*. Ph.D. dissertation, University of Amsterdam. Utrecht, Netherlands: LOT Publications.

van den Bogaerde, B., and Baker, A. E. (2005). Code mixing in mother-child interaction in Deaf families. *Sign Lang. Linguist.* 8, 153–176. doi: 10.1075/sll.8.1-2.08bog

van den Bogaerde, B., and Baker, A. E. (2009). "Bimodal language acquisition in Kodas (kids of deaf adults)," in *Hearing, Mother-Father Deaf: Hearing People in Deaf Families*, eds M. Bishop and S. L. Hicks (Washington, DC: Gallaudet University Press), 99–131.

Volterra, V., and Taeschner, T. (1978). The acquisition and development of language by bilingual children. *J. Child Lang.* 5, 311–326. doi: 10.1017/S0305000900007492

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Creating a communication system from scratch: gesture beats vocalization hands down

Nicolas Fay[1]*, Casey J. Lister[1], T. Mark Ellison[1] and Susan Goldin-Meadow[2]

[1] School of Psychology, University of Western Australia, Crawley, WA, Australia
[2] Department of Psychology, University of Chicago, Chicago, IL, USA

How does modality affect people's ability to create a communication system from scratch? The present study experimentally tests this question by having pairs of participants communicate a range of pre-specified items (emotions, actions, objects) over a series of trials to a partner using either non-linguistic vocalization, gesture or a combination of the two. Gesture-alone outperformed vocalization-alone, both in terms of successful communication and in terms of the creation of an inventory of sign-meaning mappings shared within a dyad (i.e., sign alignment). Combining vocalization with gesture did not improve performance beyond gesture-alone. In fact, for action items, gesture-alone was a more successful means of communication than the combined modalities. When people do not share a system for communication they can quickly create one, and gesture is the best means of doing so.

Keywords: alignment, gesture, vocalization, multimodal, signs, language origin, embodment

INTRODUCTION

And the Lord came down to see the city and the tower which the children of men builded. And the Lord said, "Behold, the people is one, and they have all one language; and this they begin to do: and now nothing will be restrained from them, which they have imagined to do. Go to, let us go down, and there confound their language, that they may not understand one another's speech." (Genesis 11:5–8, King James Version).

The Book of Genesis tells of the people of Babel, who build a tower that reaches to heaven. God, angered by their arrogance, and concerned by what the people might be capable of, imposes different unshared languages on them, reasoning that without a shared language the people would not be able to communicate, and thus not be able to successfully cooperate. This story was once used to explain the great variety of human languages (approximately 7000 different languages; Lewis, 2009).

Would confounding the language of the people of Babel have stopped them from successfully communicating with one another? This is unlikely. People have successfully established shared communication systems in the absence of a common language. This is seen in pidgins: simple languages that develop among groups who do not share a common language (Thomason and Kaufman, 1988) and in the sign languages that arise when deaf people are brought together (Kegl et al., 1999; Senghas et al., 2004). The present study seeks to determine which communication modality is best suited to establishing a shared communication system from scratch when people are prohibited from using their common language. The question of which modality is best suited to the creation of an *ad hoc* communication system can help inform one of the oldest and most controversial questions in science; the origin of language (Fitch, 2010). In the absence of direct evidence, this question cannot be answered with any certainty. But simulating a scenario in which modern humans must create a new communication system from scratch can help

us generate an informed guess. In this paper we use an experimental approach to examine which modality—non-linguistic vocalization, gesture or a combination of non-linguistic vocalization and gesture—best facilitates participants' ability to create a shared communication system with a partner. Specifically, we compare pairs of participants' communication accuracy and the extent to which they use the same signs to communicate the same meanings.

First we review the different theories of the origin of language and evidence supporting each position. Next we review experimental studies of natural spoken language and how they can be extended to deal with novel situations. We then discuss experimental-semiotic studies that examine the genesis of new communication systems when people are prohibited from using their existing language system. Finally, we state the experimental hypotheses and report the results of the present study.

VOCAL, GESTURAL, AND MULTIMODAL ACCOUNTS OF THE ORIGIN OF LANGUAGE

There are several theories of the origin of language, the most intuitively appealing being that human language developed from non-linguistic vocalizations (MacNeilage, 1998; Cheney and Seyfarth, 2005; Mithen, 2005). Vocalization is our primary means of communication, so it's easy to imagine human language evolving from the vocalizations of non-human primates. Like human speech, the vocalizations of non-human primates can be referential; vervet monkeys produce at least three predator-specific alarm calls that are understood by conspecifics (Seyfarth et al., 1980). However, anatomical and physiological constraints limit the vocal repertoire of non-human primates primarily to a small set of innately specified emotional signals. There is also evidence that non-human primates combine single calls into structurally more complex units with a different meaning, thereby expanding their vocal repertoire (Zuberbühler, 2002; Arnold and Zuberbühler,

2006). For example, when preceded by a low pitched "boom," the predator alarm calls of Campbell's monkeys are understood by another species, Diana monkeys, to indicate a lower level of direct threat than when the alarm calls are not preceded by a boom (Zuberbühler, 2002). Combinatorial patterning of this kind may have acted as a precursor to syntax. Cheney and Seyfarth (2005) propose that these rudimentary representational abilities are exactly what we'd expect to find in a pre-linguistic ancestor.

This view is challenged by a competing explanation; that language originated through gesture (Hewes, 1973; Corballis, 2003; Arbib, 2005). The brief timeframe in which some new sign languages have become established supports a gesture-first account (Kegl et al., 1999; Sandler et al., 2005). Several other phenomena point to the naturalness of gesture: people of all cultures gesture while they speak (Feyereisen and de Lannoy, 1991); blind people gesture (Iverson and Goldin-Meadow, 1998); speaking adults can successfully adopt gesture as their sole means of communication at the request of experimenters (Goldin-Meadow et al., 1996) or when the environment dictates (e.g., when working in a noisy sawmill; Meissner and Philpott, 1975); many of the lexical items that hearing children produce in the earliest stages of language learning appear first in gesture and only later move to the verbal lexicon (Iverson and Goldin-Meadow, 2005); young deaf children whose profound hearing losses prevent them from acquiring spoken language, and whose hearing parents have not exposed them to sign language, turn to gesture to communicate, and fashion a system of signs, called *homesign*, that contains the fundamental properties of human language (Goldin-Meadow, 2003). Perhaps the most compelling evidence in favor of a gesture-first account is that attempts to teach non-human primates to talk have failed (Hayes, 1952), whereas attempts to teach them a gestural language have been moderately successful (Gardner and Gardner, 1969; Savage-Rumbaugh et al., 1986). This, in addition to the greater flexibility of ape gestures (compared to vocal signals; Pollick and de Waal, 2007), suggests our closest relative is better equipped to communicate by gesture than by speech.

A multimodal account assumes that the earliest forms of language were not restricted to a single modality. Instead, communication occurred by any means available. Bickerton dubs this the "catch-as-catch-can" evolution of language (Bickerton, 2007, p. 512), in which language evolved from whatever rudimentary gestures or sounds were able to communicate meaning effectively. In support of this position it has been observed that, during conversation, bilinguals in a spoken and a signed language often blend their communication across the different modalities (Emmorey et al., 2008), and hearing children produce their first two-element "sentences" in gesture + speech combinations (point at bird + "nap") and only later produce them entirely in speech ("bird nap") (Iverson and Goldin-Meadow, 2005; Özçalışkan and Goldin-Meadow, 2005). Thus, given the opportunity, people use both modalities. Perniss et al. (2010) argue for a multimodal account, pointing out that vocalization-only and gesture-only explanations for language origin are both burdened with explaining why the other form of communication also exists and how it arose. They argue that the neural systems controlling vocalization and gesture are so tightly integrated because these systems have been connected from the beginning (see also Goldin-Meadow and McNeill, 1999).

EXPERIMENTAL STUDIES: EXTENDING SPOKEN LANGUAGE

Acts of reference, in which individuals refer to an object, emotion, action or some other specifiable thing, are ubiquitous to everyday communication. Several tasks have been developed to experimentally examine the referential use of language. In these tasks the experimenter assigns the participants' communicative intentions, whether this involves describing an object or giving directions to a location (for a review see Krauss and Fussell, 1996).

By having participants describe objects that lack a pre-existing name, researchers have examined the process through which people establish joint reference. One participant, the director, communicates a series of abstract shapes from an array to a partner, the matcher, who tries to identify each shape from their array. Interacting partners extend their linguistic system by creating new labels for these novel shapes (e.g., Krauss and Weinheimer, 1964; Clark and Wilkes-Gibbs, 1986). Furthermore, participants' shape descriptions, which initially are elaborate, become increasingly succinct and abstract, such that a shape first described as "Looks like a Martini glass with legs on each side" is referred to as "Martini" over the course of successive references (Krauss and Fussell, 1996, p. 679). Thus, once a shared label has been mutually agreed upon, or grounded, directors use more efficient descriptions that are understood by the matcher. Similar refinement is seen in speech-accompanying gestures (Hoetjes et al., 2011). Interaction is crucial to this process; without it, the referring expressions are longer and more complex (Krauss and Weinheimer, 1966; Hupet and Chantraine, 1992).

Other referential communication tasks show that participants' referring expressions become shared, or aligned, through interaction. Garrod and Anderson (1987) examined the linguistic descriptions used by pairs of participants working together to navigate through a computerized maze. Unlike the shape description task where participant role is typically fixed as either director or matcher, in the maze game both participants give and receive location descriptions (i.e., there is role-switching). Garrod and Anderson (1987) observed that, as the task progressed, pairs of interacting participants increasingly used the same description schemes to communicate locations on the maze. For example, if one participant used a coordinate scheme to communicate a maze location (e.g., "I'm in position A4") their partner was disproportionately likely to use the same spatial description scheme. Similar interactive alignment is observed for other aspects of linguistic form, including syntax (Branigan et al., 2000) and prosody (Giles et al., 1992). This incremental coupling between production and comprehension processes can explain why conversation is easy: linguistic representations activated by the speaker prime similar representations in the listener, and these representations retain enough activation such that when it is the listener's turn to speak they are reused (and readily understood by the previous speaker; Garrod and Pickering, 2004).

Together, these studies show that language can be rapidly extended to deal with novel situations. They demonstrate that interaction is critical for efficient communication, and that when people alternate speaker and listener roles, they increasingly share, or align upon, the same communication system. Experimental-semiotic studies adopt similar experimental paradigms to study the process through which new communication systems arise and

evolve when participants are denied use of their existing linguistic system.

EXPERIMENTAL STUDIES: CREATING NEW COMMUNICATION SYSTEMS

Because language does not leave fossils, it is difficult to test theories of the origin of language. Moreover, because observational studies of the emergence of pidgins and new sign languages lack experimental control, it is difficult to confidently isolate the variables critical to the genesis and evolution of new languages. Experimental-semiotic studies try to overcome these problems by studying the emergence of new communication systems under controlled laboratory conditions. They do this by creating a situation where participants must communicate without using their existing language system (for a review see Galantucci and Garrod, 2011). Typically, participants communicate in a novel modality, for example, through drawing (Galantucci, 2005; Garrod et al., 2007), through gesture (Goldin-Meadow et al., 1996; Gershkoff-Stowe and Goldin-Medow, 2002; Goldin-Meadow et al., 2008; Langus and Nespor, 2010; Fay et al., 2013) or movement (Scott-Phillips et al., 2009; Stolk et al., 2013), and the experimenters study how communication systems evolve across repeated interactions between the human agents.
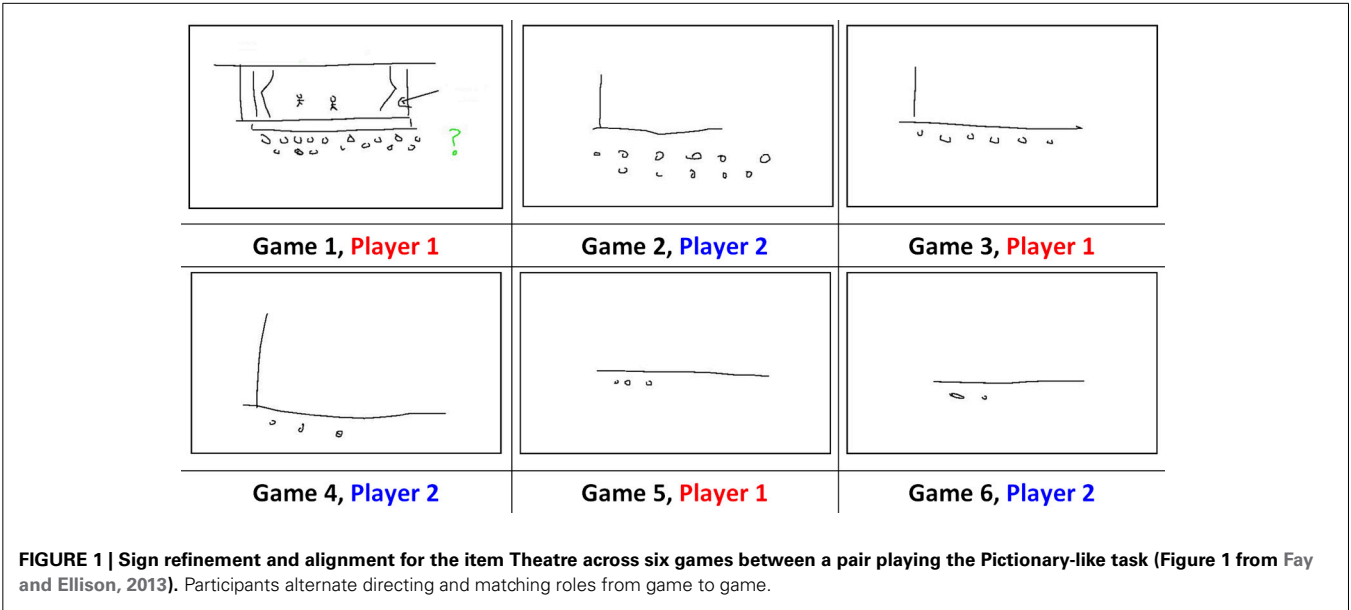
A key finding of relevance to the present study is that participants initially use iconic signs to ground shared meanings, and over subsequent interactions these signs become increasingly aligned, symbolic and language-like (Garrod et al., 2007; Fay et al., 2010; Garrod et al., 2010). In Garrod et al. (2007) participants communicated a set of recurring items to a partner by drawing on a shared whiteboard (e.g., *Art Gallery*, *Drama*, *Theatre*). Much like the game Pictionary™, participants were not allowed to speak or use numbers or letters in their drawings. This procedure forced them to create a new communication system from scratch. As participants repeatedly played the game, the form of their signs changed: for example, at game 1 the sign used to communicate Theater was a visually complex iconic drawing of a theater, including a stage, curtains, actors and an audience, whereas by game 6 it had evolved into a simple symbolic drawing, communicated by a line and two circles. Notice also that the signs produced by each member of the pair became increasingly similar, or aligned over games (see **Figure 1**). Like spoken referential communication studies, sign refinement is only seen when participants interact with a partner. Repeated drawing without interaction does not lead to such abstraction (in fact, the drawings become more complex; Garrod et al., 2007, 2010).

Experimental-semiotic studies indicate that, when people are prohibited from using their existing language, they use iconic signs to ground shared meanings. Once grounded, the signs become increasingly simplified and aligned, much like spoken language referential communication studies. This process makes the signs easier to execute and comprehend. Given that gesture lends itself more naturally to the production of iconic signs than vocalization, Fay et al. (2013) reasoned that gesture has the potential to provide a superior modality for bootstrapping a communication system from scratch. They tested this prediction in a referential communication study where pairs of participants communicated sets of items (Emotions, Actions, Objects) using non-linguistic vocalization, gesture, or a combination of non-linguistic vocalization and gesture. As predicted, gesture proved more effective (more communication success) and more efficient (less time taken) than vocalization at communicating the different items. Combining gesture with vocalization did not improve performance beyond gesture alone. This finding suggests an important role for gesture in the origin of the earliest human communication systems.

PRESENT STUDY

Communication is not possible unless people share a common inventory of sign-meaning mappings. The present study tests the extent to which communication modality drives the

FIGURE 1 | Sign refinement and alignment for the item Theatre across six games between a pair playing the Pictionary-like task (Figure 1 from Fay and Ellison, 2013). Participants alternate directing and matching roles from game to game.

creation of such an inventory. As in Fay et al. (2013), pairs of participants were assigned to a communication modality (non-linguistic vocalization, gesture, non-linguistic vocalization and gesture combined) and tried to communicate a set of recurring items (Emotions, Actions, Objects) to their partner. Sign alignment was not possible in the Fay et al. (2013) study because participants were allocated to fixed roles (director or matcher) for the duration of the experiment. In the present study participants alternate roles from game to game, allowing them to copy (or not) features of their partners' signs. This simple change in design lets us determine the extent to which partners align their signs.

Our first hypothesis is that communication success will be higher for gesture than for non-linguistic vocalization. Such a result would confirm the findings reported by Fay et al. (2013). Our second hypothesis speaks to the affordance offered by combining modalities. If combining modalities is advantageous because the two modalities offer independent sources of information, we would expect communication success to be higher in the combined modality compared to gesture-alone. While no difference in communication success between gesture and the combined modality was reported by Fay et al. (2013) this may be due to a lack of statistical power. The present study uses almost twice as many participants and double the number of communication games.

The main focus of this paper is alignment. Intuitively, people must establish a mutually shared sign-to-meaning mapping before they can align their sign systems. The extent to which sign-to-meaning mappings are shared is indexed by communication success. Following our first hypothesis (greater communication success in the gestural modality), we therefore expect greater agreement in sign-to-meaning mappings in the gestural modality. Agreement in interpretation, while not enforcing alignment, i.e., use of the same meaning-to-sign mapping, is a prerequisite for the latter. Thus, our third hypothesis is that there will be greater alignment in the gestural modality than in the vocalization modality. Based on our prediction that communication success will be highest in the combined modality, our fourth hypothesis is that alignment will be strongest when both modalities are used.

Our final hypothesis concerns the relationship between communication success and alignment. As discussed above, communication success can be seen as an index of sign-to-meaning agreement, which enables alignment. Evidence of this is seen in a study that established a link between linguistic alignment and performance on a joint cooperative task (Fusaroli et al., 2012). Hypothesis five is that there will be a positive correlation between communication success and sign alignment in each modality.

METHODS

This study received approval from the University of Western Australia Ethics Committee. All participants viewed an information sheet before giving written consent to take part in the study. The information sheet and consent form were both approved by the aforementioned Ethics Committee.

PARTICIPANTS

Ninety-two undergraduate psychology students (57 females) participated in exchange for partial course credit or payment.

Participants were tested in unacquainted pairs, in testing sessions lasting 1 h. All were free of any visual, speech or hearing impairment.

TASK AND PROCEDURE

Participants completed the task in pairs. Participants were randomly assigned to the role of director or matcher and switched roles at the end of each game, e.g., Participant 1 was the director on Game 1 and Participant 2 was the matcher; on Game 2 Participant 2 was the director and Participant 1 was the matcher, and so on across Games 1–12. Each game consisted of 18 trials. On any trial, the director's task was to communicate a specific item from an ordered list of 24 items (18 target items and 6 distractor items presented on a sheet of A4 paper) that were known to both participants. Items were drawn from three categories (Emotion, Action, Object) and included easily confusable items such as *Tired* and *Sleeping* (see **Table 1** for a complete listing of the experimental items). The director's task was to communicate the first 18 items from their list in the given order. On the director's list the first 18 items were always the target items (presented in a different random order on each game). The 18 target items were the same on each game and for each pair of participants. On the director's list the final 6 items were always the distracter items (presented in a different random order on each game). The 6-distractor items were the same on each game and for each pair of participants. Distractor items were included to ensure that matchers could not use a process of elimination to identify the target items. The distracter items were never communicated. The matcher's list was presented in a different random order on a sheet of A4 paper (with all 24 items presented in a different random order). The matcher's task was to indicate the order in which each item was communicated by inserting the trial number beside the relevant item. Participants played the game 12 times with the same partner, using the same item set on each game (i.e., each participant directed 6 times and matched 6 times).

Each pair was randomly allocated to one of three communication modalities: Vocal ($N = 28$), Gesture ($N = 28$) or Combined (gesture plus vocalization) ($N = 26$). In each modality, participants were seated at opposite sides of a round **Table 1** meter in diameter. Those in the Vocal modality were told they could make

Table 1 | The experimental items directors tried to communicate to matchers (distracter items are given in italic).

Emotion	Action	Object
Tired	Fleeing	Rock
Pain	Sleeping	Fruit
Angry	Fighting	Predator
Hungry	Throwing	Water
Disgust	Chasing	Tree
Danger	Washing	Hole
Happy	*Eating*	*Mud*
Ill	*Hitting*	*Rain*

Target and distracter items were fixed across conditions and throughout the experiment.

any sounds, and as many sounds (including vocal repetitions) as they wished, but were not permitted to use words. In this modality, participants sat back-to-back, ruling out the use of visual signals. Once the director had communicated each of the 18 target items, the pair swapped roles and the next game began. The new director then communicated the same 18 target items, but in a different random order. This process was repeated until 12 games had been played. Those in the Gesture modality faced one another across the table. All communication was limited to gesture (hand, body and face) and vocalizing was prohibited. Participants were permitted to make any gestures, and as many gestures (including gesture repetitions) as they wished. Participants in the Combined modality followed the same procedure as those in the Gesture modality, but were permitted to vocalize in addition to gesturing. In each modality, matchers indicated to directors they had made their selection by saying "ok," and then privately inserting the trial number (1–18) next to the selected item. Matchers were only permitted to select an item once.

Irrespective of role, both participants could interact within a trial (e.g., a matcher might seek clarification by frowning or by grunting). As in most human communication studies, participants were not given explicit feedback with regard to their communication success (e.g., Clark and Wilkes-Gibbs, 1986; Garrod and Anderson, 1987; Anderson et al., 1991; Garrod et al., 2007). All communication was recorded using a pair of digital video cameras (one trained on each participant).

RESULTS

We took two measures of the developing communication systems: effectiveness and alignment. Effectiveness was operationalized as the percentage of items successfully identified by the matcher. Alignment measured the degree to which participants used the same signs as their partner for the same items.

EFFECTIVENESS

Effectiveness measures how successful the signs were at identifying their referent. As **Figure 2** shows, participants' identification success improved across games 1–12 in all modalities and for each item type (Emotion, Action and Object). In the Gesture and Combined modalities, the different item types were communicated with similar success. In the Vocal modality, Emotion items were more successfully communicated than Action items (in the early games but not in the late games) and Action items were more successfully communicated than Object items (across all Games). Communication effectiveness was very high (and close to ceiling) in the Gesture and Combined modalities, and much lower in the Vocal modality.

For simplicity, and to reduce between-game variance, the factor Games was collapsed into three bins corresponding to Early (1–4), Middle (5–8), and Late (9–12) Games. Participants' mean percent accuracy scores were entered into a mixed design ANOVA that treated Modality (Vocal, Gesture, Combined) as a between-participant factor and Item (Emotion, Action, Object) and Game (Early, Middle, Late) as within. All main effects were significant, as were each of the two-way interactions and the three-way Modality-by-Item-by-Game interaction (see **Table 2A**).

To understand the 3-way interaction we ran three separate Item-by-Game ANOVAs for each level of Modality (Vocal, Gesture, Combined). The 3-way interaction can be explained by the Item-by-Game interaction in the Vocal modality, and the sole main effect of Game in the Gesture and Combined modalities (**Tables 2B–D**, respectively). Although communication success improved across games for each item type in each modality, in the
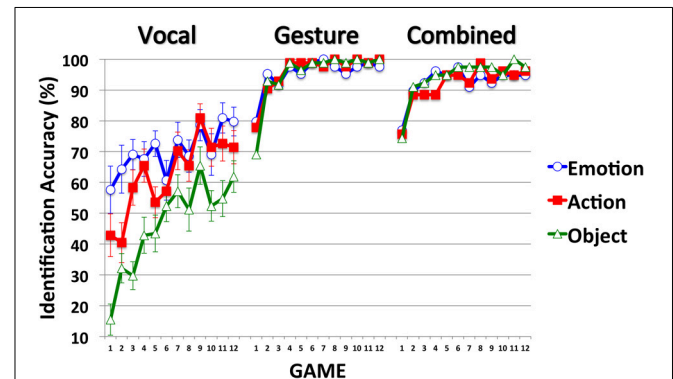

FIGURE 2 | Mean identification accuracy across Items (Emotion, Action, Object) and Games (1–12), expressed as percentage scores, for participants in the Vocal, Gesture, and Combined modalities. Error bars indicate the standard errors of the means (included only on items in the Vocal modality to reduce unnecessary clutter).

Table 2 | (A) Results of the 3 × 3 × 3 ANOVA that treated Modality (Vocal, Gesture, Combined) as a between-participant factor and Item (Emotion, Action, Object) and Game (Early, Middle, Late) as within-participant factors. Results of the 3 × 3 ANOVAs for each level of Modality: (B) Vocal, (C) Gesture, and (D) Combined.

	df	F	p	Partial eta squared
(A) OVERALL 3 × 3 × 3 ANOVA				
Modality	2, 38	123.02	<0.001	0.87
Item	2, 76	22.17	<0.001	0.37
Game	2, 76	70.64	<0.001	0.65
Modality × item	4, 76	25.93	<0.001	0.58
Modality × game	4, 76	8.46	<0.001	0.31
Item × game	4, 152	6.52	<0.001	0.15
Modality × item × game	8, 152	2.42	=0.017	0.11
(B) VOCAL				
Item	2, 26	39.53	<0.001	0.75
Game	2, 26	28.55	<0.001	0.69
Item × game	4, 52	5.27	<0.001	0.29
(C) GESTURE				
Item	2, 26	0.11	=0.90	0.01
Game	2, 26	48.98	<0.001	0.79
Item × game	4, 52	1.60	=0.19	0.11
(D) COMBINED				
Item	2, 24	0.45	=0.64	0.04
Game	2, 24	25.40	<0.001	0.68
Item × game	4, 48	0.96	=0.44	0.07

Vocal modality the different items were communicated with different levels of success. In the Early games, Emotion items were more successfully communicated than Action items, and Action items were more successfully communicated than Object items. By the late games, Emotion and Action items were communicated with equal success, and both were communicated with greater success than Object items. In contrast, the different item types were communicated with similar success in both the Gesture and Combined modalities.

In support of Hypothesis 1, and as observed by Fay et al. (2013), communication success was higher for each item type in the Gesture and Combined modalities than in the Vocal modality: Emotion [$Fs_{(1, 26/25)} > 28.12$, $ps < 0.001$, $\eta_p^2 s > 0.53$], Action [$Fs_{(1, 26/25)} > 65.54$, $ps < 0.001$, $\eta_p^2 s > 0.72$] and Object items [$Fs_{(1, 26/25)} > 226.23$, $ps < 0.001$, $\eta_p^2 s > 0.90$]. Hypothesis 2, which predicted higher communication success in the Combined modality, was not supported. Communication success was comparable across the Gesture and Combined modalities for Emotion and Object items [$Fs_{(1, 25)} < 1.09$, $ps < 0.31$, $\eta_p^2 s < 0.04$]. However, Gesture proved more successful than the Combined modality at communicating Action items [$Fs_{(1, 25)} = 4.84$, $ps = 0.037$, $\eta_p^2 s = 0.16$]. Thus, with more statistical power, the null effect reported by Fay et al. (2013) reached statistical significance in the present study.

Gesture is a more effective means of communication than vocalization, and combining gesture with vocalization does not improve communication success beyond gesture alone. In fact, it may make it worse.

ALIGNMENT

An illustrative example of communication from a pair of participants in the Gesture modality, sampled from the early (1–4) and late games (9–12) is given in **Figure 3**. Initially a variety of different signs were used to communicate the object "predator." Eventually the partners aligned on the same simplified sign.

A bespoke coding scheme was developed to elucidate the process through which pairs of participants establish a shared communication system. The coding scheme was designed to assess sign variation and the extent to which pairs of participants were able to negotiate a stable and shared sign for each meaning over the course of the experiment. Broadly, we predict that sign stability/sharedness will increase across games in each modality. The coding scheme was applied to the signs produced by directors in each modality, as they communicated the 18 different target items across games 1–12. Each sign was coded into one of the following categories: Innovate (new, previously unseen sign for this item), Copy (replication of partner's sign for the same item from the immediately prior game), Copy and Simplify (simplified version of partner's sign for the same item from the immediately prior game), Copy and Elaborate (more complex version of partner's sign for the same item from the immediately prior game), Reuse Self (participant reuses a sign for the same item from their prior turn as director), and Throwback (participant uses a sign for the same item from an earlier game, but not one from their partner's immediately prior turn as director, or from their own immediately prior turn as director). The changing frequencies of the different sign categories are shown in **Figure 4** (collapsed across

the different item types). Video examples from each modality are available at http://comlab.me/ComLab/GestureBeatsVocal.html.

Innovation is the only option at Game 1 as there are no earlier signs to copy. Hence, there is 100% sign Innovation at Game 1 in each modality. From this point onwards, sign Innovation decreases dramatically across games. This decrease in Innovation is most strongly observed in the Gesture and Combined modalities, compared to the Vocal modality. As Innovation decreases, sign Copying increases over games. Sign Copying is more strongly observed in the Gesture and Combined modalities (78 and 71% respectively by Game 12) compared to the Vocal modality (52%). Sign Copy and Simplify was prominent at Game 2 in the Gesture and Combined modalities (18 and 20%, respectively) and was almost absent by Game 12 (<1%). Copy and Elaborate was less frequent but showed a similar pattern (10 and 13%, respectively, at Game 2 and <1% by Game 12). Sign Copy was less frequent in the Vocal modality (52% at Game 12), as was Copy and Simplify (4% at Game 2) and Copy and Elaborate (5% at Game 2). Participants in the Vocal modality frequently Reused the sign they produced on their prior turn as director (42% at Game 12, compared to 21 and 23% in the Gesture and Combined modalities). Throwbacks were too infrequent to compare (occurring on only 1.2% of trials across Games 2–12). The more frequent sign Copying observed in the Gesture and Combined modalities indicates that the signs were more shared, or aligned, in these modalities, compared to the Vocal modality.

We tested this observation by comparing the overall frequency of Sign Copying (by combining the Copy, Copy and Simplify and Copy and Elaborate categories) across the different modalities. Game 1 was not included in the analysis as sign Copying was not possible. As **Figure 5** shows, sign copying increased across games in each modality, and for each item type. Sign copying is comparable across modalities for Emotion items, but is higher in the Gesture and Combined modalities for Action and Object items.

The factor Game was again collapsed into three bins corresponding to Early (2–4), Middle (5–8), and Late (9–12) Games. Participants' mean percent Copying scores were entered into a mixed design ANOVA that treated Modality (Vocal, Gesture, Combined) as a between-participants factor and Item (Emotion, Action, Object) and Game (Early, Middle, Late) as within. This returned main effects for Modality, Item and Game [$Fs_{(2, 38/76)} < 6.41$, $ps < 0.003$, $\eta_p^2 s > 0.14$]. There was also a Modality-by-Item and Modality-by-Game interaction [$Fs_{(2, 76)} < 4.90$, $ps < 0.001$, $\eta_p^2 s > 0.21$]. No other effects reached statistical significance [$Fs < 2.08$, $ps > 0.09$, $\eta_p^2 s < 0.05$].

As **Figure 5** shows, sign alignment in the Vocal modality mirrors identification accuracy: stronger alignment on Emotion items followed by Action and Object items. A different pattern is observed in the Gesture and Combined modalities where stronger alignment is seen for Action items followed by Objects and Emotion items. More importantly, pairwise comparisons indicate a similar level of alignment for Emotion items across the different modalities [$ts_{(26/25)} < 1.44$, $ps > 0.16$, $ds < 0.542$], but stronger alignment for Action and Object items in the Gesture and Combined modalities compared to the Vocal modality [$ts_{(26/25)} > 4.55$, $ps < 0.001$, $ds > 1.75$]. A similar level of alignment was observed for each item type in the Gesture and

FIGURE 3 | Signs used by a pair in the Gesture modality to communicate the object "predator" at Games 1–4 (Early) and 9–12 (Late). Game number is given in the leftmost column. At Game 1 Director A claws at the air (correctly identified by partner). At Game 2 Director B mimes a hulking movement, with her arms out to the side. Next she throws her arms up in fright before miming a running action (incorrectly identified). At Game 3 Director A copies Director B; she throws her arms in the air and mimes walking like a hulk (incorrectly identified). At Game 4 Director B points over her shoulder, mimes walking like a hulk, then mimes running (correctly identified). Communication is simple, aligned and successful from Game 9: both partners communicate "predator" by raising their arms in their air to mime a hulk walking.
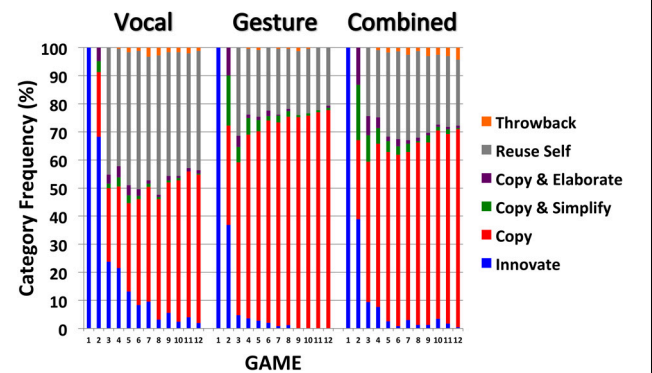
FIGURE 4 | Mean frequency, expressed as percentage scores, of Innovate, Copy, Copy, and Simplify, Copy and Elaborate, Reuse Self and Throwback signs across Games 2–12 for participants in the Vocal, Gesture and Combined modalities. Error bars indicate the standard errors of the means.
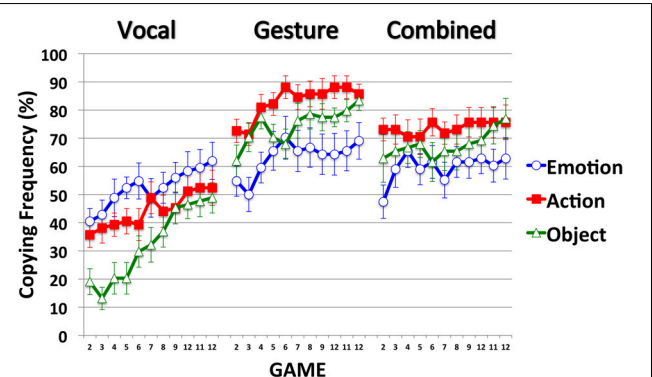

FIGURE 5 | Mean copying frequency, expressed as percentage scores, of signs across Items (Emotion, Action, Object) and Games (2–12) for participants in the Vocal, Gesture and Combined modalities. Error bars indicate the standard errors of the means.

Combined modalities [$ts_{(25)} < 1.69$, $ps > 0.10$, $ds < 0.65$]. Thus, the Modality-by-Item interaction can be explained by a similar level of alignment across modalities for Emotion items, and stronger alignment for Action and Object items in the Gesture and Combined modalities (compared to the Vocal modality).

The Modality-by-Game interaction is explained by the strong increase in sign copying across games in the Vocal modality [$F_{(2, 26)} = 22.82$, $p < 0.001$, $\eta_p^2 = 0.64$] and Gesture modality [$F_{(2, 26)} = 13.17$, $p < 0.001$, $\eta_p^2 = 0.50$] and the weaker, marginal increase in sign copying in the Combined modality [$F_{(2, 24)} = 2.95$, $p = 0.057$, $\eta_p^2 = 0.21$]. Pairwise comparisons indicate that sign alignment is stronger for Early, Middle and Late games in the Gesture and Combined modalities, compared to the Vocal modality [$ts_{(26/25)} > 2.69$, $ps < 0.013$, $ds > 1.04$]. Sign alignment scores were similar in the Gesture and Combined modalities [$ts_{(25)} < 1.74$, $ps > 0.094$, $ds < 0.67$].

In summary, there was greater sign alignment when participants could use gesture to communicate. This finding supports Hypothesis 3. Hypothesis 4, that sign alignment will be stronger in the Combined modality, was not supported. In fact, sign alignment increased more strongly in the Gesture modality compared to the Combined modality.

EFFECTIVENESS AND ALIGNMENT

To what extent are communication effectiveness and sign alignment linked? Hypothesis 5 predicts a positive correlation between the two. This would be consistent with communication success promoting sign alignment and/or sign alignment promoting communication success. To determine if a relationship exists, participants' mean identification accuracy scores (collapsed across games 2–12) were correlated with their mean copying scores (collapsed across games 2–12). A strong positive correlation was observed in the Vocal $[r_{(14)} = 0.81, p_{one-tailed} < 0.001]$ and Combined modalities $[r_{(13)} = 0.75, p_{one-tailed} = 0.001]$, and a moderate correlation was observed in the Gesture modality $[r_{(14)} = 0.45, p_{one-tailed} = 0.055]$. The correlations in the Gesture and Combined modalities are all the more remarkable given the lack of variation in identification accuracy scores (due to the near ceiling effect; see **Figure 6**). This pattern supports Hypothesis 5.

DISCUSSION

The present study experimentally tested the influence of modality (vocal, gesture, or a combination of the two) on how people establish a shared communication system from scratch when they cannot use an existing language system. Gesture proved to be a more effective means of communication than non-linguistic vocalization, supporting Hypothesis 1[1]. Hypothesis 2, that combining the two modalities would prove more effective than gesture alone, was not supported. In fact, Gesture was comparable to the Combined modality for Emotion and Object items, and was more successful at communicating Action items.

The primary motivation behind the present study was to test how modality affects the establishment of a shared inventory of signs. This shared inventory arises via progressive sign alignment (Pickering and Garrod, 2004). Gesture enabled stronger sign alignment than Vocalization for Action and Object items, but not for Emotion items, partly supporting Hypothesis 3. Hypothesis 4, that combining the two modalities would produce stronger alignment than gesture alone, was not supported. In fact, the increase in sign alignment across games was stronger

---

[1]Gesture might be more effective (communication success) than vocalization because vocalization suffers greater interference from participants' first (spoken) language. This is possible, although it is equally possible that communication success in the vocalization-only condition was facilitated by participants' first (spoken) language. An issue for an interference explanation is that the different item types (emotion, action, object) showed a differential pattern of communication success. General interference from an already established vocal language would predict a similar performance decrement in the vocalization-only modality for the different item types relative to the gesture modality. Further research with deaf signers or bimodal bilinguals (e.g., English-ASL) is needed to make a definite determination about whether performance on the task is affected by participants' existing language system.

for Gesture alone than for the Combined modality. Hypothesis 5 predicted a positive correlation between communication success and sign alignment. Consistent with a link between linguistic alignment and task performance (Fusaroli et al., 2012), a positive correlation between communication success and sign alignment was returned for each modality. Of course, causality cannot be determined: communication success may promote sign alignment or sign alignment may promote communication success, or both. We suspect causality acts in both directions.

WHY ARE COMMUNICATION SUCCESS AND SIGN ALIGNMENT HIGHER FOR GESTURE THAN FOR VOCALIZATION?

Among modern day humans, with modern brains and mastery of at least one spoken language, the present study demonstrates the superiority of gesture over non-linguistic vocalization as a solution to the Babel problem. In this context gesture is a more precise modality of communication than non-linguistic vocalization. We believe this precision arises from its greater affordance of motivated signs: iconic signs that communicate through resemblance, or indexical signs that communicative via a natural association between sign and referent. For Vocalization, the link between sign and referent tends to be arbitrary, that is, symbolic, with the exception of a small inventory of onomatopoeic and sound-symbolic expressions (see Shintel and Nusbaum, 2007). For example, participants in the Gesture modality could close their eyes and pretend sleep to communicate *Tired* (a natural index of tiredness), clench their fist and pantomime throwing a punch to communicate *Fighting* (an iconic representation) or peel an imaginary banana to communicate *Fruit* (an indexical representation). These motivated relationships between sign and referent are much less obvious for Vocalization. They do exist for some Emotion items, for example, making yawn noises to communicate *Tired* (a vocal index of tiredness), but are mostly absent for Action and Object items. For instance, it's hard to imagine a motivated vocalization that could be used to communicate *Chasing* or *Mud*. Our data support this: in the Combined modality, vocalization was added to gesture on
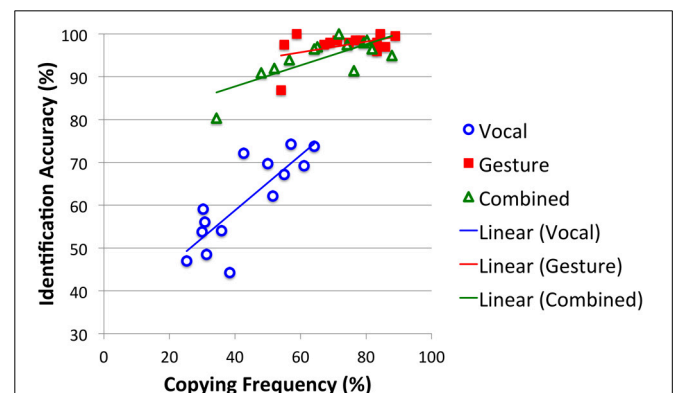

FIGURE 6 | Correlation between Identification Accuracy (mean percent of items correctly identified across games 2–12) and Sign Copying (mean percent of signs copied across games 2–12) for participants in the Vocal, Gesture and Combined modalities.

54% of trials for Emotion items, 26% of trials for Object items and 14% of trials for Action items (and remained stable across games).

Our study suggests that affordances of motivated signs are essential to bootstrapping a set of shared sign-meaning mappings when people cannot draw on a pre-existing inventory of shared conventional signs. Once the sign-meaning mappings have been grounded, interlocutors can reduce the complexity of the signs—causing them to evolve into more symbol-like forms (Garrod et al., 2007, 2010)—and align their signs. Both processes reduce the cost of sign production and comprehension (Pickering and Garrod, 2004, 2013). These local interactive processes underpin the propagation of a shared inventory of conventional signs in larger populations, as shown in computer simulations (Steels, 2003; Barr, 2004; Tamariz et al., under review), natural spoken language studies (Garrod and Doherty, 1994), experimental semiotic studies (Fay et al., 2008, 2010; Fay and Ellison, 2013) and naturalistic studies of recently formed sign languages (Goldin-Meadow et al., under review; Kegl et al., 1999).

Returning to theories of the origin of language, our results suggest a strong role for gesture due to its affordance of motivated signs. In the absence of a conventional language, it is unlikely that our ancestors would have passed up the opportunity to use motivated signs, in particular gesture, to get their point across. This is to not to rule out a multimodal, "catch-as-catch-can" account (Bickerton, 2007, p. 512), far from it: when permitted, participants often used vocalization in combination with gesture, especially for Emotion items (54% of trials in the Combined modality). The productive use of vocalization as an index of emotions (see also Sauter et al., 2010) fits with our position that motivated signs are likely to have played an important role in establishing the earliest human communication systems. However, it is important to be clear that in the present study vocalization played a supporting role, always occurring in the company of gesture and not replacing gesture. Gesture, we propose, played the primary role in bootstrapping the earliest human communication systems on account of its affordance of motivated signs. Today, the vocal modality is primary and gesture plays a supporting role. The dynamics of the rise of predominantly vocal language, and the reasons for it, are targets for future research (see Goldin-Meadow and McNeill, 1999; Corballis, 2002; Corballis, for some suggestions such as the affordance of vocalization for communication in the dark).

WHY IS GESTURE BETTER THAN GESTURE PLUS VOCALIZATION AT COMMUNICATING ACTION ITEMS?

The finding that Gesture alone was more successful at communicating Action items than the Combined modality warrants further consideration. One candidate explanation is that participants were distracted by the auditory information conveyed in the Vocal modality (Spence et al., 2000). This explanation is plausible because Vocal-only communication is less precise than Gesture-only communication in the present study. If information conveyed in the vocal channel acts as a distractor from information conveyed in the visual channel, we would expect a negative correlation between vocalization frequency and communication

success. That is, more frequent vocalization will be associated with lower communication success. Participants' mean vocalization frequency (percent of trials in which vocalization occurred in addition to gesture collapsed across games 1–12) was correlated with their mean communication success. A moderate negative correlation was returned [$r_{(13)} = -0.39$, $B = -0.138$, $p_{\text{one-tailed}} = 0.095$], indicating that more frequent vocalization is associated with lower communication success for Action items. Although a similar negative correlation was observed for Object items [$r_{(13)} = -0.48$, $B = -0.075$, $p_{\text{one-tailed}} = 0.045$], its gradient is shallower compared to that of Action items, meaning that the negative impact of vocalization on communication success was less strongly felt. The correlation for Emotion items did not approach statistical significance [$r_{(13)} = -0.13$, $B = -0.030$, $p_{\text{one-tailed}} = 0.339$].

Why did vocalization negatively impact communication success for Action items? More than Object or Emotion items, Action items offer an opportunity for embodiment in the Gesture modality (Lakoff and Johnson, 1999; Hostetter and Alibali, 2008). By taking a character viewpoint, participants can simulate the action as the sign: to communicate *Throwing* the participant can extend their right arm back and mime the throwing of a ball. Embodied action is less direct for Emotion items, which are internal states, and Object items, which have no direct human role to take (although some participants pantomimed a human interaction with the object). The infrequent addition of vocalization when communicating Action items in the Combined modality (14% of trials) reflects the intrinsic fit between gesture and actions. This fit is reinforced by Action items exhibiting the strongest levels of sign alignment in the Gesture modality, compared to the other item types (see **Figure 5**). Against this natural fit between gesture and actions, supplementary vocalizations distract the matcher from a channel that is ideally suited to the communication of actions.

EXPERIMENTAL GESTURE CREATION COMPARED TO NATURALISTIC GESTURE CREATION

Our study has some limitations, the most important of which is that our participants have modern day brains and already speak a language. The second is that our participants are creating labels out of context, which is not likely to be the way language emerges on the ground. Finally, we ask our participants to create words, but we do not ask them to string those words together, that is, to create sentences. Studies of naturalistic language creation in homesigners address some, but not all, of these limitations. As mentioned earlier, homesigners are individuals whose profound hearing losses prevent them from acquiring the spoken language that surrounds them, even when given intensive instruction in speech. They are, in addition, born to hearing parents who do not expose them to a conventional sign language. Under these circumstances, we might expect that a homesigner would not communicate at all. But homesigners do communicate, and they use gesture to do so (Goldin-Meadow, 2003).

Homesigners thus do not have usable input from a conventional language model and are truly creating language from scratch (although they do have modern day brains). Moreover, the gestures homesigners create are all used in a naturalistic

context. Like the participants in our study, young homesigners use iconic gestures to refer to actions. However, they prefer to use pointing gestures, rather than iconic gestures, to refer to objects (they rarely refer to emotions, but neither do young children learning conventional language). Over time, homesigners use iconic gestures more and more often to refer to objects as well as actions, and they develop morphological devices to distinguish between the two uses (Goldin-Meadow et al., 1994). Not surprisingly, because they are communicating with hearing individuals who do not share their gesture systems, homesigners rarely produce gestures whose forms are not transparently related to their referents; that is, they rarely produce non-iconic gestures. For the same reason, their gestures do not lose their iconicity over time. Nevertheless, these iconic gestures are combined with other gestures to form structured sentences. Homesigners combine their pointing gestures (and later their iconic gestures referring to objects) with iconic gestures referring to actions, and use these gesture sentences to communicate about the here-and-now and the non-present, to make generic statements, to tell stories, to talk to themselves, and even to refer to their own gestures—that is, to serve the central functions of language (Goldin-Meadow, 2003). The fact that homesigners begin the process of language creation by using gesture to convey actions fits nicely with our finding that gesture affords an easily accessible way to convey action, and suggests that our experimental paradigm is capturing an early stage of an important aspect of language creation.

In addition to creating gestures in a naturalistic context, homesigners also differ from our participants in that they are interacting with hearing individuals who have no interest in creating a shared gesture system with them. Homesigners in the U.S. are typically born to hearing parents who would like their deaf children to learn to speak; they therefore often do not learn sign language themselves and rarely gesture to their children without talking at the same time (Flaherty and Goldin-Meadow, 2010). The gestures they produce are thus co-speech gestures, which are qualitatively different in form from homesign (Goldin-Meadow et al., 1996). In other words, the homesigners' parents do *not* align their gestures with their children's gestures (Goldin-Meadow and Mylander, 1983). Interestingly, although homesigners display many of the grammatical features of natural language in their gestures, their gestures do not form a stable lexicon in the same way that our participants' gestures do. Goldin-Meadow et al. (under review) studied adult homesigners in Nicaragua and found that they used different gestures from each other to label the same object, which is not surprising given that the homesigners did not know one another. More importantly from our point of view, each individual homesigner used a variety of gestures to label a single object and was not consistent within him or herself. The homesign data thus support the conclusions from our study—that alignment between speakers is essential for a lexicon to stabilize.

CONCLUSION

The Tower of Babel story asks if people can communicate when they do not share a common language. The present study experimentally tests the affordances offered by vocalization and gesture when creating a common inventory of signs from scratch. Gesture outperformed non-linguistic vocalization both in terms

of communication success and in terms of the creation of a common inventory of sign-meaning mappings. Combining vocalization with gesture did not improve performance beyond gesture alone; in fact, it sometimes proved deleterious. We argue that the benefit of gesture lies in its ability to communicate through motivated signs, and this makes it an excellent modality for language creation.

REFERENCES

Anderson, A. H., Clark, A., and Mullin, J. (1991). Introducing information in dialogues: forms of introduction chosen by young speakers and the responses elicited from young listeners. *J. Child Lang.* 18, 663–687. doi: 10.1017/S0305000900011302

Arbib, M. (2005). From monkey-like action recognition to a human language: an evolutionary framework for neurolinguistics. *Behav. Brain Sci.* 28, 105–167. doi: 10.1017/S0140525X05000038

Arnold, K., and Zuberbühler, K. (2006). Language evolution: semantic combinations in primate calls. *Nature* 441, 303. doi: 10.1038/441303a

Barr, D. J. (2004). Establishing conventional communication systems: is common knowledge necessary? *Cogn. Sci.* 28, 937–962. doi: 10.1207/s15516709cog2806_3

Bickerton, D. (2007). Language evolution: a brief guide for linguists. *Lingua* 117, 510–526. doi: 10.1016/j.lingua.2005.02.006

Branigan, H. P., Pickering, M. J., and Cleland, A. A. (2000). Syntactic co-ordination in dialogue. *Cognition* 75, B13–B25. doi: 10.1016/S0010-0277(99)00081-5

Cheney, D. L., and Seyfarth, R. M. (2005). Constraints and preadaptations in the earliest stages of language evolution. *Ling. Rev.* 22, 135–159. doi: 10.1515/tlir.2005.22.2-4.135

Clark, H. H., and Wilkes-Gibbs, D. (1986). Referring as a collaborative process. *Cognition* 22, 1–39. doi: 10.1016/0010-0277(86)90010-7

Corballis, M. C. (2002). *From Hand to Mouth: The Origins of Language*. Princeton, Oxford: Princeton University Press.

Corballis, M. C. (2003). From mouth to hand: gesture, speech, and the evolution of right-handedness. *Behav. Brain Sci.* 26, 199–260. doi: 10.1017/S0140525X03000062

Emmorey, K., Borinstein, H. B., Thompson, R., and Gollan, T. H. (2008). Bimodal bilingualism. *Biling (Camb. Engl).* 11, 43.

Fay, N., Arbib, M., and Garrod, S. (2013). How to bootstrap a human communication system. *Cogn. Sci.* 37, 1356–1367. doi: 10.1111/cogs.12048

Fay, N., and Ellison, T. M. (2013). The cultural evolution of human communication systems in different sized populations: usability trumps learnability. *PLoS ONE* 8:e71781. doi: 10.1371/journal.pone.0071781

Fay, N., Garrod, S., and Roberts, L. (2008). The fitness and functionality of culturally evolved communication systems. *Philos. Trans. R. Soc. B Biol. Sci.* 363, 3553–3561. doi: 10.1098/rstb.2008.0130

Fay, N., Garrod, S., Roberts, L., and Swoboda, N. (2010). The interactive evolution of human communication systems. *Cogn. Sci.* 34, 351–386. doi: 10.1111/j.1551-6709.2009.01090.x

Feyereisen, P., and de Lannoy, J. (1991). *Gestures and Speech: Psychological Investigations*. New York, NY: Cambridge University Press.

Fitch, W. (2010). *The Evolution of Language*. Cambridge, UK: Cambridge University Press. doi: 10.1017/CBO9780511817779

Flaherty, M., and Goldin-Meadow, S. (2010). "Does input matter? Gesture and homesign in Nicaragua, China, Turkey, and the USA," in *Paper presented at the Proceedings of the Eighth Evolution of Language Conference* (Utrecht). doi: 10.1142/9789814295222_0066

Fusaroli, R., Bahrami, B., Olsen, K., Roepstorff, A., Rees, G., Frith, C., et al. (2012). Coming to terms: quantifying the benefits of linguistic coordination. *Psychol. Sci.* 23, 931–939. doi: 10.1177/0956797612436816

Galantucci, B. (2005). An experimental study of the emergence of human communication systems. *Cogn. Sci.* 29, 737–767. doi: 10.1207/s15516709cog0000_34

Galantucci, B., and Garrod, S. (2011). Experimental semiotics: a review. *Front. Hum. Neurosci.* 5:11. doi: 10.3389/fnhum.2011.00011

Gardner, R., and Gardner, B. (1969). Teaching sign langue to a chimpanzee. *Science* 165, 664–672. doi: 10.1126/science.165.3894.664

Garrod, S., and Anderson, A. (1987). Saying what you mean in dialogue: a study in conceptual and semantic co-ordination. *Cognition* 27, 181–218. doi: 10.1016/0010-0277(87)90018-7

Garrod, S., and Doherty, G. (1994). Conversation, coordination and convention - an empirical investigation of how groups establish linguistic conventions. *Cognition* 53, 181–215. doi: 10.1016/0010-0277(94)90048-5

Garrod, S., Fay, N., Lee, J., Oberlander, J., and MacLeod, T. (2007). Foundations of representation: where might graphical symbol systems come from? *Cogn. Sci.* 31, 961–987. doi: 10.1080/03640210701703659

Garrod, S., Fay, N., Rogers, S., Walker, B., and Swoboda, N. (2010). Can iterated learning explain the emergence of graphical symbols? *Interact. Stud.* 11, 33–50. doi: 10.1075/is.11.1.04gar

Garrod, S., and Pickering, M. J. (2004). Why is conversation so easy? *Trends Cogn. Sci.* 8, 8–11. doi: 10.1016/j.tics.2003.10.016

Gershkoff-Stowe, L., and Goldin-Medow, S. (2002). Is there a natural order for expressing semantic relations? *Cogn. Psychol.* 45, 375–412. doi: 10.1016/S0010-0285(02)00502-9

Giles, H., Coupland, N., and Coupland, J. (1992). "Accommodation theory: communication, context and consequences," in *Contexts of Accommodation*, eds H. Giles, J. Coupland, and N. Coupland (Cambridge: Cambridge University Press), 1–68.

Goldin-Meadow, S. (2003). *The Resilience of Language: What Gesture Creation in Deaf Children Can Tell Us About How Children Learn Language*. New York, NY: Psychology Press.

Goldin-Meadow, S., Butcher, C., Mylander, C., and Dodge, M. (1994). Nouns and verbs in a self-styled gesture system: what's in a name? *Cogn. Psychol.* 27, 259–319. doi: 10.1006/cogp.1994.1018

Goldin-Meadow, S., and McNeill, D. (1999). "The role of gesture and mimetic representation in making language the province of speech," in *The Descent of Mind: Psychological Perspectives on Hominid Evolution*, eds M. C. Corballis and S. E. G. Lea (New York, NY: Oxford University Press), 155–172.

Goldin-Meadow, S., McNeill, D., and Singleton, J. (1996). Silence is liberating: removing the handcuffs on grammatical expression in the manual modality. *Psychol. Rev.* 103, 34–55. doi: 10.1037/0033-295X.103.1.34

Goldin-Meadow, S., and Mylander, C. (1983). Gestural communication in deaf children: noneffect of parental input on language development. *Science* 221, 372–374. doi: 10.1126/science.6867713

Goldin-Meadow, S., So, W. C., Özyürek, A., and Mylander, C. (2008). The natural order of events: how speakers of different languages represent events nonverbally. *Proc. Natl. Acad. Sci. U.S.A.* 105, 9163–9168. doi: 10.1073/pnas.0710060105

Hayes, C. (1952). *The Ape in Our House*. London: Gollancz.

Hewes, G. W. (1973). Primate communication and the gestural origin of language. *Curr. Anthropol.* 14, 5–24. doi: 10.1086/201401

Hoetjes, M., Koolen, R., Goudbeek, M., Krahmer, E., and Swerts, M. (2011). "GREEBLES greeble greeb: on reduction in speech and gesture in repeated references," in *Paper Presented at the Proceedings of the 33rd Annual Conference of the Cognitive Science Society*, Boston, MA.

Hostetter, A. B., and Alibali, M. W. (2008). Visible embodiment: gestures as simulated action. *Psychon. Bull. Rev.* 15, 495–514. doi: 10.3758/PBR.15.3.495

Hupet, M., and Chantraine, Y. (1992). Changes in repeated references: collaboration or repetition effects? *J. Psycholinguist. Res.* 21, 485–496.

Iverson, J. M., and Goldin-Meadow, S. (1998). Why people gesture when they speak. *Nature* 396, 228. doi: 10.1038/24300

Iverson, J. M., and Goldin-Meadow, S. (2005). Gesture paves the way for language development. *Psychol. Sci.* 16, 367–371. doi: 10.1111/j.0956-7976.2005.01542.x

Kegl, J., Senghas, A., and Coppola, M. (1999). "Creation through contact: sign language emergence and sign language change in Nicaragua," in *Language Creation and Language Change: Creolization, Diachrony and Development*, ed M. DeGraff (Cambridge, MA: MIT Press), 179–237.

Krauss, R. M., and Fussell, S. R. (1996). "Social psychological models of interpersonal communication," in *Social Psychology: Handbook of basic principles*, eds E. T. Higgins and A. W. Kruglanski (New York, NY: Guilford Press), 655–701.

Krauss, R. M., and Weinheimer, S. (1964). Changes in the length of reference phrases as a function of social interaction: a preliminary study. *Psychon. Sci.* 1, 113–114. doi: 10.3758/BF03342817

Krauss, R. M., and Weinheimer, S. (1966). Concurrent feedback, confirmation and the encoding of referents in verbal communication. *J. Pers. Soc. Psychol.* 4, 343–346. doi: 10.1037/h0023705

Lakoff, G., and Johnson, M. (1999). *Philosophy in the Flesh: The Embodied Mind and its Challenge to Western Thought*. New York, NY: Basic books.

Langus, A., and Nespor, M. (2010). Cognitive systems struggling for word order. *Cogn. Psychol.* 60, 291–318. doi: 10.1016/j.cogpsych.2010.01.004

Lewis, M. P. (2009). *Ethnologue: Languages of the World Sixteenth Edition*. Dallas, TX: SIL International. Available online at: Online version: http://www.ethnologue.com

MacNeilage, P. F. (1998). The frame/content theory of evolution of speech production. *Behav. Brain Sci.* 21, 499–546. doi: 10.1017/S0140525X98001265

Meissner, M., and Philpott, S. (1975). The sign language of sawmill workers in British Columbia. *Sign Lang. Stud.* 9, 291–308. doi: 10.1353/sls.1975.0010

Mithen, S. (2005). *The Singing Neanderthals: The Origins of Music, Language, Mind and Body*. London: Weidenfeld and Nicolson.

Özçalışkan, Ş., and Goldin-Meadow, S. (2005). Gesture is at the cutting edge of early language development. *Cognition* 96, B101–B113. doi: 10.1016/j.cognition.2005.01.001

Perniss, P., Thompson, R., and Vigliocco, G. (2010). Iconicity as a general property of language: evidence from spoken and signed languages. *Front. Psychol.* 1:277. doi: 10.3389/fpsyg.2010.00227

Pickering, M. J., and Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behav. Brain Sci.* 27, 169–226. doi: 10.1017/S0140525X04000056

Pickering, M. J., and Garrod, S. (2013). An integrated theory of language production and comprehension. *Behav. Brain Sci.* 36, 329–347. doi: 10.1017/S0140525X12001495

Pollick, A. S., and de Waal, F. B. M. (2007). Ape gestures and language evolution. *Proc. Natl. Acad. Sci. U.S.A.* 104, 8184–8189. doi: 10.1073/pnas.0702624104

Sandler, W., Meir, I., Padden, C., and Aronoff, M. (2005). The emergence of grammar: systematic structure in a new language. *Proc. Natl. Acad. Sci. U.S.A.* 102, 2661–2665. doi: 10.1073/pnas.0405448102

Sauter, D. A., Eisner, F., Ekman, P., and Scott, S. K. (2010). Cross-cultural recognition of basic emotions through nonverbal emotional vocalizations. *Proc. Natl. Acad. Sci. U.S.A.* 107, 2408–2412. doi: 10.1073/pnas.0908239106

Savage-Rumbaugh, S., McDonald, K., Sevcik, R., Hopkins, W., and Rubert, E. (1986). Spontaneous symbol acquisition and communicative use by pygmy chimpanzees (*Pan paniscus*). *J. Exp. Psychol. Gen.* 115, 211–235. doi: 10.1037/0096-3445.115.3.211

Scott-Phillips, T. C., Kirby, S., and Ritchie, G. R. S. (2009). Signalling signalhood and the emergence of communication. *Cognition* 113, 226–233. doi: 10.1016/j.cognition.2009.08.009

Senghas, A., Kita, S., and Ozyurek, A. (2004). Children creating core properties of language: evidence from an emerging sign language in Nicaragua. *Science* 305, 1779–1782. doi: 10.1126/science.1100199

Seyfarth, R. M., Cheney, D. L., and Marler, P. (1980). Monkey responses to three different alarm calls: evidence of predator classification and semantic communication. *Science* 210, 801–803. doi: 10.1126/science.7433999

Shintel, H., and Nusbaum, H. C. (2007). The sound of motion in spoken language: visual information conveyed by acoustic properties of speech. *Cognition* 105, 681–690. doi: 10.1016/j.cognition.2006.11.005

Spence, C., Ranson, J., and Driver, J. (2000). Cross-modal selective attention: on the difficulty of ignoring sounds at the locus of visual attention. *Percept. Psychophys.* 62, 410–424. doi: 10.3758/BF03205560

Steels, L. (2003). Evolving grounded communication for robots. *Trends Cogn. Sci.* 7, 308–312. doi: 10.1016/S1364-6613(03)00129-3

Stolk, A., Verhagen, L., Schoffelen, J.-M., Oostenveld, R., Blokpoel, M., Hagoort, P., et al. (2013). Neural mechanisms of communicative innovation. *Proc. Natl. Acad. Sci. U.S.A.* 110, 14574–14579. doi: 10.1073/pnas.1303170110

Thomason, S. G., and Kaufman, T. (1988). *Language Contact, Creolization, and Genetic Linguistics*. Berkley, CA: University of California Press.

Zuberbühler, K. (2002). A syntactic rule in forest monkey communication. *Anim. Behav.* 63, 293–299. doi: 10.1006/anbe.2001.1914

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Moving from hand to mouth: echo phonology and the origins of language

Bencie Woll *

Deafness, Cognition and Language Research Centre, University College London, London, UK

*Correspondence:
Bencie Woll, Deafness, Cognition and Language Research Centre, University College London, 49 Gordon Square, London WC1H 0PD, UK
e-mail: b.woll@ucl.ac.uk

Although the sign languages in use today are full human languages, certain of the features they share with gestures have been suggested to provide information about possible origins of human language. These features include sharing common articulators with gestures, and exhibiting substantial iconicity in comparison to spoken languages. If human proto-language was gestural, the question remains of how a highly iconic manual communication system might have been transformed into a primarily vocal communication system in which the links between symbol and referent are for the most part arbitrary. The hypothesis presented here focuses on a class of signs which exhibit: "echo phonology," a repertoire of mouth actions which are characterized by "echoing" on the mouth certain of the articulatory actions of the hands. The basic features of echo phonology are introduced, and discussed in relation to various types of data. Echo phonology provides naturalistic examples of a possible mechanism accounting for part of the evolution of language, with evidence both of the transfer of manual actions to oral ones and the conversion of units of an iconic manual communication system into a largely arbitrary vocal communication system.

Keywords: sign language, echo phonology, language origins, neuroscience of sign language, mouth gestures

INTRODUCTION

In the past 50 years, the study of how human language evolved (evolutionary linguistics) has again become a prominent feature of linguistic discourse. A complete theory of language evolution is beyond the scope of this paper, including as it must, consideration of brain function and anatomical changes in the vocal tract. We are concerned here with only one part of the process—the previously hypothesized shift from a primarily gestural or vocal-gestural communication system to spoken language (see section Historical Perspectives below) and how such a shift could have provided a mechanism for converting iconic manual symbols into arbitrary vocal symbols. Data from the sign languages of Deaf[1] communities will provide an insight into this mechanism.

Since home signing (gesture systems) can appear in the absence of linguistic input (Goldin-Meadow, 2003), sign languages used by Deaf communities have sometimes been regarded as primitive in comparison to spoken languages, and as representing earlier forms of human communication. However, linguistic research over the past 40 years has demonstrated that sign languages are in fact full natural languages with complex grammars (Stokoe, 1960; Klima and Bellugi, 1979; Sutton-Spence and Woll, 1999). The creators and users of all known sign languages are humans with "language-ready brains." Nevertheless, it is possible that sign share features in common with evolutionary precursors of spoken language.

These features include sharing common articulators with non-linguistic communication (i.e., gestures), and exhibiting substantial iconicity in comparison to spoken languages. This iconicity is present in signs representing abstract concepts as well as in those that represent concrete objects and actions. The form of many signs [examples from British Sign Language (BSL)] depict part or all of a referent or an action associated with a referent, such as EAT, PAINT (holding and using a paintbrush), CAT (whiskers), BIRD (beak). Signs referring to cognitive activities (THINK, UNDERSTAND, KNOW, LEARN, etc.) are generally located at the forehead, while signs relating to emotional activities (FEEL, INTERESTED, EXCITED, ANGRY) are located on the chest and abdomen; signs with the index and middle fingers of the hand extended and separated ("V" handshape) relate to concepts of "two-ness": TWO, BOTH, TWO-OF-US, WALK (legs), LOOK, READ (eyes). The pervasiveness of iconicity (even where heavily conventionalized) is striking, in both sign languages and gestures.

If human proto-language was gestural or vocal-gestural, the question remains as to how such a communication system with a high degree of iconicity might link to the development of articulated words in spoken language, in which the links between symbol and referent are, for the most part, seen as arbitrary. Posing the question in this way, and regarding sign languages as "manual" ignores the rich and complex role played by other articulators: body, face, and, in particular, the mouth. As well as the actions performed by the hands, sign languages also make use of mouth actions of various types. The theory proposed here relates to one subgroup of mouth actions: "echo phonology" (Woll and Sieratzki, 1998; Woll, 2001). These are a

---

[1]"Deaf" with an upper-case "D" is used to refer to membership of a sign language-using community and includes both hearing and (audiologically) deaf individuals.

set of mouth actions unrelated to spoken language, and which occur obligatorily in a number of sign languages alongside certain manual signs. They are characterized by "echoing" on the mouth certain of the articulatory activities of the hands.

Three data sources are discussed here: narratives in 3 different European sign languages, anecdotal observations of hearing individuals bilingual in BSL and English, and functional imaging studies with deaf signers. These provide evidence of a possible mechanism in the evolution of spoken language by which iconic symbols in a manual communication system could have converted into a vocal communication system with arbitrary links between symbol and referent.

HISTORICAL PERSPECTIVES

Many writers have suggested that human vocal language may have evolved from manual gestures. What is required to sustain such a claim is a plausible mechanism by which primarily manual actions could have transformed themselves into vocal actions. One mechanism (not even requiring communicative gesturing as an intermediate stage) was suggested by Darwin in *The Expression of Emotions in Man and Animals* (1872):

> "there are other actions [of the mouth] which are commonly performed under certain circumstances…and which seem to be due to imitation or some sort of sympathy. Thus, persons cutting anything may be seen to move their jaws simultaneously with the blades of the scissors. Children learning to write often twist about their tongues as their fingers move, in a ridiculous fashion." (Darwin, 1872, p. 34)

Henry Sweet (1888) extended this notion to encompass a transition from manual gesture to "lingual gesture":

> "Gesture.. helped to develop the power of forming sounds while at the same time helping to lay the foundation of language proper. When men first expressed the idea of "teeth," "eat," "bite," it was by pointing to their teeth. If the interlocutor's back was turned, a cry for attention was necessary which would naturally assume the form of the clearest and most open vowel. A sympathetic lingual gesture would then accompany the hand gesture which later would be dropped as superfluous so that ADA or more emphatically ATA would mean "teeth" or "tooth" and "bite" or "eat," these different meanings being only gradually differentiated." (Sweet, 1888, pp. 50–52)

To Sweet, therefore, should go the credit for hypothesizing that a "lingual gesture accompanying a natural hand gesture" could be a key link between gesture and spoken language. However, he provides no evidence for such a process, failing to explain more general features of what he calls sympathetic lingual gestures.

Richard Paget (1930) attempted to find evidence for such a theory. Like Sweet, Paget claimed that the earliest human language was a language of gestures, in which manual actions were unconsciously copied by movements -of the mouth, tongue, or lips.

> "Originally man expressed his ideas by gesture, but as he gesticulated with his hands, his tongue, lips and jaw unconsciously followed suit…The consequence was that when, owing to pressure of other business, the principal actors (the hands) retired from

the stage…their understudies—the tongue, lips, and jaw—were already proficient in the pantomimic art." (Paget, 1930, p. 133)

He supplies a number of examples of this process:

> "Another…example may be given, namely, in connection with the beckoning gesture—commonly made by extending the hand, palm up, drawing it inwards toward the face and at the same time bending the fingers inwards toward the palm. This gesture may be imitated with the tongue, by protruding, withdrawing, and bending up its tip as it re-enters the mouth.
>
> If this "gesture" be blown or voiced, we get a resultant whispered or phonated *word*, like **ɛdə**, **ɛdə**, or **ɛdra**…suggestive of …our English word "hither"." (Paget, 1930, p. 138)

Paget's theory (known as the "ta-ta" theory from another example suggesting parallels between waving goodbye and flapping the tongue) was developed further by Swadesh (1971). He provides another example of its application:

> "…a word like the Latin *capio*, I take, or English *capture*, whose root begins with a *k* sound and ends in the sound *p*, made by closing the lips. It has been suggested that the formation of the *k* sound at the back of the mouth, while the lips are open, is comparable to the open hand. The closing of the lips, then, is analogous to the fingers closing with the thumb as one takes hold of an object. Thus the pronunciation of the root *capio* is like the action of taking. Of course not all words are to be explained in this way; in fact, only a few. And yet the possibility that some words developed in this way is not denied by other qualities also evident in language." (Swadesh, 1971, p. 4)

Paget's theory can only be validated if there is evidence for a historical process by which manual gestures were reflected in movements of the lips and tongue, which were in turn associated with the production of speech-sounds. One weakness of the approach of Paget and the others is that they all suggest that the mouth actions share underlying imagery with the associated iconically-motivated manual gesture, leaving open the question of how a hypothesized highly iconic manual communication system could have subsequently led to spoken language, with its generally arbitrary links between symbol and referent.

Hewes (1973) serves as a point of connection between the writers of the late nineteenth and early twentieth century and contemporary writings on language evolution. Kendon (2010) in a review of Fitch (2010) summarizes Hewes' view that primate gestures served as a better point of comparison with human language than their vocalizations. Hewes did recognize, however that a challenge for a gestural origin of human language was the need to account for the switch from manual to vocal communication. His suggested reasons included the greater convenience of speaking (it could be used in the dark and while the hands were occupied), and an increase in vocabulary and ease of lexical retrieval. He also supported Paget's (1930) hypothesis, discussed above.

Recent studies (Erhard et al., 1996; Rizzolatti and Arbib, 1998; Rizzolatti and Craighero, 2004) provide such evidence of links between brain areas associated with language and areas controlling movement of the hands and arms (also see below). However, such findings have not been used to suggest a mechanism in

language evolution for the twin shifts from hand to mouth and from iconic to arbitrary symbols.

CONTEMPORARY EVIDENCE
NEUROBIOLOGICAL PERSPECTIVES

Studies of neurons in the monkey brain by Rizzolatti and colleagues since 1996 (Rizzolatti et al., 1996; Rizzolatti and Craighero, 2004) have identified "mirror neurons," which fire when a primate observes another individual (monkeys and humans) making specific reaching and grasping movements. The mirror system, in temporal, parietal, and frontal regions, is part of a system specialized for perceiving and understanding biological motion. Although research has not shown a mapping of vocalization production onto perception of vocalizations, this mapping is implicit in Liberman and Mattingly's (1985) motor theory of speech perception, which proposes that speech is understood in terms of its articulation, rather than its perception. It should also be noted that the anatomical proximity of neurons in the pre-motor cortex relating to hand and mouth functions may relate evolutionarily to the involvement of both in activities such as eating. The relationships between mouth actions related to eating, and mouth actions found in spoken language, have been discussed in detail by MacNeilage (1998). Meguerditchian and Vauclair (2008), describe shared features in the co-occurrence of manual and vocal gestures in non-human primates.

In a series of studies, Gentilucci and colleagues have shown that mouth actions are related to manual actions. When participants were asked to grasp objects of different sizes while articulating syllables such as /ba/ there was a parallel increase in the mouth opening and voice spectra of syllables pronounced simultaneously. Semantically congruent words and gestures also show interaction effects not seen in incongruent pairings (Gentilucci, 2003; Gentilucci and Corballis, 2006; Barbieri et al., 2009). Bernardis and Gentilucci (2006), describing the relationship of words and emblems in processing and execution, hypothesize that a system relating actions to syllables might have evolved into a system relating symbolic gestures to words, and importantly, draw on neurological evidence about the role of Broca's area in both gesture and language.

GESTURE AND SPEECH

A number of theorists have postulated that manual gesture (on its own, without consideration of vocalization or mouth gesture) is the origin of language. Rizzolatti and Arbib (1998) align with the earlier nineteenth and twentieth century writers, seeing gesture as fading once speech has emerged:

> "Manual gestures progressively lost their importance, whereas, by contrast, vocalization acquired autonomy, until the relation between gestural and vocal communication inverted and gesture became purely an accessory factor to sound communication" (Rizzolatti and Arbib, 1998, p. 193).

In such models, gesture is seen as unintegrated with speech—both in modern human communication and in human evolution.

McNeill et al. (2008) provides a strong set of arguments against this position. They argue that a unimodal communication system,

using gesture or sign alone, could not have evolved into modern human communication, which is primarily bimodal (gesture and speech). They suggest that if such a phase existed, it was not a proto-language, but a precursor of mimicry and pantomime. They argue that a "gesture-first" theory:

> "incorrectly predicts that speech would have supplanted gesture, and fails to predict that speech and gesture became a single system. It is thus a hypothesis about the origin of language that almost uniquely meets Popper's requirement of falsifiability—and *is* falsified, doubly so in fact. (McNeill et al., 2008, p. 12)"

Another thread in the "supplantation of gesture by speech" argument relates to the advantages of speech over gesture (Corballis, 2003). McNeill et al. (2008) have argued that speech is the default form of human communication because it has fewer dimensions, is more linear, is non-imagistic (and hence more arbitrary, with the potential for a larger lexicon), etc. Given this asymmetry, McNeill and colleagues argue that even though speech and gesture are selected jointly, it would still be the case that speech is the medium of linguistic segmentation:

> "Sign languages—their existence as full linguistic systems—impresses many as a reason for gesture-first, but in fact, historically and over the world, manual languages are found only when speech is unavailable; the discrete semiotic then transferring to the hands. As we shall see later, this transfer takes place automatically. So it is not that gesture is incapable of carrying a linguistic semiotic, it is that speech (to visually disposed creatures) does not carry the imagery semiotic." (McNeill et al., 2008, p. 13)

HANDS AND MOUTH IN SIGN LANGUAGE
MOUTH ACTIONS AND OTHER NON-MANUAL ARTICULATORS

As mentioned above, sign languages of the deaf offer a unique perspective on language, since they embody the structural and communicative properties of spoken language, while existing entirely within a wholly visual-gestural medium. Among other insights, they enable investigators to clarify the core components of language in distinction to those that reflect input or action characteristics of the language system. This difference is reflected in the articulators on which languages in the two modalities rely. Sign languages use both manual and non-manual articulators, including the head, face and body (e.g., Liddell, 1978; Sutton-Spence and Woll, 1999). Within the face, eye actions such as eye narrowing, changes in direction of gaze and eyebrow actions (raise/lower) play important roles in SL communication (Crasborn, 2006). In addition, although sign languages are not historically related to the spoken languages of their surrounding hearing communities, sign languages do borrow elements from spoken language (Sutton-Spence and Woll, 1999). Thus some mouth actions (mouthings) are derived from spoken language, while other mouth actions (mouth gestures) are unrelated to spoken languages (see **Figure 1** below).

In a study of narratives in three European sign languages (Crasborn et al., 2008) mouth actions were found throughout (**Table 1**). There is striking uniformity in the percentage of signs accompanied by mouth gestures (35–39%), with greater variation

across the three languages in the percentage of signs accompanied by mouthings (26–51%).

Mouthings

Sign languages can borrow mouth actions from spoken words—speech-like actions accompanying manual signs that can disambiguate manually homonymous forms. These are considered to be borrowings, rather than contact forms reflecting bilingualism in a spoken and signed language, since there is evidence that signers can learn these without knowing the source spoken language. These serve to disambiguate "manual homonyms": signs with similar or identical manual forms. For example, the BSL signs ASIAN and BLUE, are manually identical (see **Figure 3C** below). To distinguish which meaning is meant, mouthings are incorporated, derived from the mouth actions used when speaking the words "Asian" or "blue."

Adverbials

Adverbials are arrangements of the mouth which are used to signal manner and degree (e.g., to indicate that an action is performed with difficulty or with ease; to indicate if an object is very small or very large, etc.). In **Enaction** (sometimes called mouth-for-mouth), the action performed by the mouth represents that action directly (e.g., in CHEW, the mouth performs a "chewing" action, while the sign is articulated on the hands).

ECHO PHONOLOGY

The term **Echo Phonology** (Woll and Sieratzki, 1998; Woll, 2001, 2009) is used for a class of mouth actions that are obligatory in the citation forms of lexical signs. In the BSL sign TRUE (see **Figure 3D** below), the upper hand moves downwards to contact the lower hand, and this action is accompanied by mouth closure, synchronized with the hand contact. This category of mouth gesture differs from adverbial mouth arrangements as the mouth gesture forms part of the citation form of the manual

FIGURE 1 | Mouth actions in sign language.

sign, and unlike adverbial mouth gestures, do not carry additional meaning. Crasborn et al. (2008) refer to this category of mouth gestures as "semantically empty." Signs with echo phonology appear incomplete or ill-formed in their citation form if the mouth gesture is not present.

The term "echo phonology" is used, since the mouth action is a visual and motoric "echo" of the hand action in a number of respects: onset and offset, dynamic characteristics (speed and acceleration) and type of movement (e.g., opening or closing of the hand, wiggling of the fingers). Echo phonology mouth gestures are not derived from or related in any other way to mouth actions representing spoken words; in the citation form of these signs they are an obligatory component, and are presumably constrained by the common motor control mechanisms for hands and mouth discussed above. The citation forms of the signs in which they are found require the presence of the mouth gesture to be well-formed, and the mouth gesture always includes some movement such as inhalation or exhalation, or a change in mouth configuration (opening or closing) during the articulation of the sign: for example, BSL signs EXIST (wiggling of fingers, no path movement, accompanied by [∫∫]); TRUE (active hand makes abrupt contact with palm of passive hand, accompanied by [am]—see **Figure 3D** below); DISAPPEAR (spread hands close to "flat o" shape, accompanied by [θp]).

The essential dependence of the mouth gesture on the articulatory features of the manual movement can be seen in three BSL signs all meaning "succeed" or "win." Three different oral patterns of mouthing co-occur with these signs, and one cannot be substituted for the other. In SUCCEED, the thumbs are initially in contact, but move apart abruptly as the mouth articulates [pa]. In WIN, the hand rotates at the wrist repeatedly as the mouth articulates [hy]; and in WON, the hand closes to a flat O, while the mouth articulates [ʌp]. Most importantly, the action of the mouth in signs with echo phonology, while echoing that of the hands, is not in itself iconic.

The parallel movements of the hands and mouth found in echo phonology can also be seen in the production of the BSL sign DISAPPEAR (**Figure 2**). Both hands are open and the tongue is protruding at the onset of the manual sign. The notation tiers show that during the movement of the sign, as the hands close, the tongue retracts.

SYLLABLES OCCURRING IN ECHO PHONOLOGY IN BSL

The following elements (**Table 2**) have been identified, although it is likely that this is not an exhaustive list. It is not known what inventories exist in other sign languages. Some articulatory features are given for them; and since echo phonology is a feature of a language used by deaf people, no voiced-voiceless distinction is
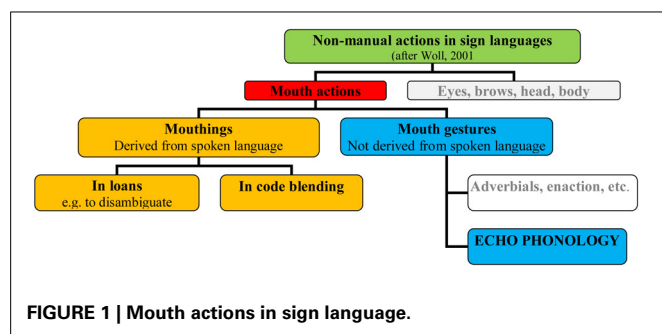
Table 1 | Comparison of hand/mouth actions in three sign languages.

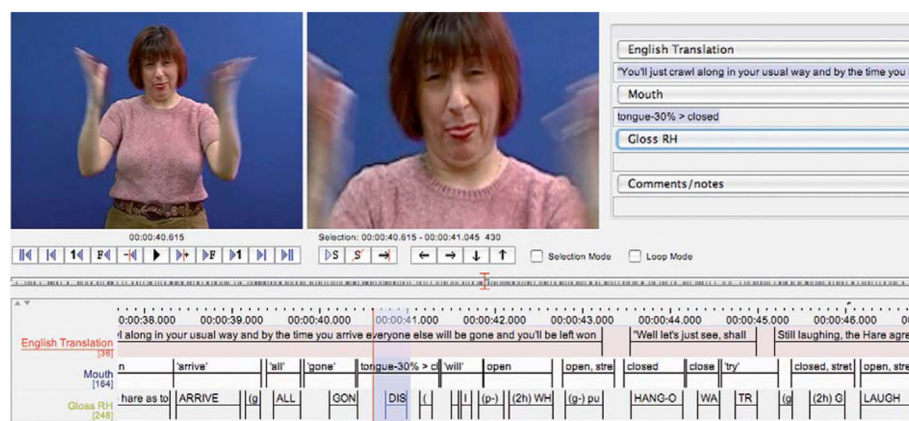| Language	Number of Manual signs	Number of mouth gestures	% of signs accompanied by mouth gestures	Number of mouthings	% of signs accompanied by mouthings
British Sign Language	1552	539	35	560	36
Sign Language of the Netherlands	1162	458	39	299	26
Swedish Sign Language	1619	624	39	831	51

FIGURE 2 | BSL sign disappear showing initial configuration of mouth.

Table 2 | Echo phonology elements in BSL.

CONSONANTS—SYLLABLE-INITIAL	
p	Bilabial stop
f	Labio-dental stop
CONSONANTS—SYLLABLE-FINAL	
p	Bilabial stop
m	Bilabial nasal
ʕ	Glottal stop
VOCALICS	
y	Front rounded vowel
ʌ or a	Low central vowel
u	Back rounded vowel
h	Pharyngeal fricative
w	Bilabial fricative
θ	Interdental fricative
ʃ	Rounded palatal fricative
BREATH PATTERNS	
<	Exhalation
>	Inhalation

operative and almost all involve articulations at the front of the mouth or lips, where they are most visible.

The combinations of these elements result in syllables. Selected examples of signs using these syllables are given (**Table 3**).

Although echo phonology is largely voiceless in deaf signers, hearing people with deaf parents (bilinguals native in both BSL and English) frequently mix sign and speech, either in the form of code-mixing (switching between English and BSL) or—because these languages occur in different modalities (bimodal bilingualism)—by means of code blending, where elements from a spoken language appear simultaneously with elements of a sign language.

Anecdotal observations from conversations between hearing people with deaf parents (bilinguals native in both BSL and English) indicate that echo phonology appears (with or without voicing) in the form of code mixing with English in the absence

Table 3 | Examples of syllables with echo phonology.

<	done
pa	

Associated with one or two active hands, movement consists of hand separation and twisting, with single sharp action

<	<	<	not-bothered
fu	*or* fw	*or* fy	

Body contact at start of sign, movement downwards and away from body

<	exist
ʃ	

Wriggling or fingers, repeated shaking or twisting of wrists(s), no path movement

<	<	win
hw	*or* hy	

Repeated twisting of wrist, no path movement

>	thank-god
ʌp	

Closing and twisting of hand(s), sharp movement

>	>	true
am	*or* ʌm	

Hand closes and contacts passive hand, sharp movement

>	disappear (also see **Figure 2**)
θp	

Hand(s) close, sharp movement with abrupt stop

of production of the manual component. In other words, only the oral component is produced.

Examples include:

1. A: "Have you done that poster?"
   B: "[ʃʃʃ] (NOT-YET), I'll do it tomorrow" (voiceless)
2. A: "It was terrible. [ʌmp]" (END/absolutely over) (voiced)
3. A: "I couldn't get a straight answer from anyone. It was completely [pɪpɪpɪ]" (VARIED/inconsistent) (voiced)

These examples are suggestive of a possible leap from echo phonology in signs to a situation where voicing accompanies

these mouth gestures so that they begin to have independent existence as lexical items. Further research is necessary to explore whether these forms are more similar to vocal gestures or to words.

Sweet, Paget and the other early writers cited above postulated that iconicity in the mouth gesture itself was the source of spoken words. However, it is difficult to see how a mouth gesture on its own could iconically express the semantic notion of "succeed" or "true." Echo phonology illustrates a mechanism by which abstract concepts, which can be represented by iconic manual gestures, can be attached to abstract mouth gestures.

ECHO PHONOLOGY IN DIFFERENT SIGN LANGUAGES

In a study comparing narratives in three sign languages, the occurrence of echo phonology was compared with other types of mouth action. The data are drawn from the ECHO (European Cultural Heritage Online) corpus. This corpus was created as part of a European Union pilot project with the aim of demonstrating how scientific data within the humanities (including linguistics) can be made widely accessible via the Internet (Crasborn et al., 2007).

Data were collected from one male and one female Deaf native signer of each of BSL, NGT, and SSL—a total of six signers. After reading brief summaries in order to familiarize themselves with the content, signers were asked to sign to camera their own versions of five of Aesop's fables. Data were then coded with ELAN software, using a broadly defined set of transcription categories. In all, 51 min of signed material were included in this study. All annotated data from this study is freely available at the ECHO web site: http://www.let.ru.nl/sign-lang/echo.2

Echo phonology was found in all three sign languages. Of mouth gestures found in the narratives (i.e., excluding signs with mouthing), signs with echo phonology form 10.8% of mouth gestures in BSL, 12.6% in Sign Language of the Netherlands, and 16% in Swedish Sign Language (Crasborn et al., 2008).

Echo phonology has also been studied in other sign languages, including German Sign Language (Pendzich, 2013) and American Sign Language (Mather and Malkowski, 2013). Mather and Malkowski explored opening and closing movements of the mouth in detail, in particular, how mouth closing occurs when the hands contact the body, and mouth opening occurs when hand contact with the body is broken.

NEURAL CORRELATES OF ECHO PHONOLOGY

Despite the differences in the modality of the perceived signal, the neural organization of language is remarkably similar in spoken and signed language. Neuroimaging studies of native signers show similar patterns of lateralization and activation when processing spoken or signed language data. Specifically, sign language processing is associated with activation in left temporal and frontal cortex, including Broca's area (BA 44/45), just as for spoken language. (see e.g., Emmorey, 2001; Corina et al., 2003; MacSweeney et al., 2008; Newman et al., 2010 for a review). MacSweeney et al. (2002) also found no differences between BSL and English in the extent of lateralization, with both languages left-lateralized. Studies of patients with brain lesions following CVA consistently indicate that perisylvian regions of the left hemisphere support

language processing (Atkinson et al., 2005; see Woll, 2012 for a review).

Despite their similarities, the networks for spoken and sign language are not completely identical. MacSweeney et al. (2002) report that regions which showed more activation for BSL than audiovisual English included the middle occipital gyri, bilaterally, and the left inferior parietal lobule (BA 40). In contrast, audiovisual English sentences elicited greater activation in superior temporal regions than BSL sentences (pp. 1589–1590).

With these considerations in mind, Capek et al. (2008) explored the sensitivity of the cortical circuits used for language processing to the specific articulators used, not only comparing speech and signing but examining activation during perception of signs with English mouthing, with echo phonology, and with no mouth actions. In their fMRI experiment, lists of lexical items were presented to deaf native signers. These comprised: (1) silently articulated English words with no hand action (SR); (2) BSL signs with no mouth action (hands only—Man); (3) BSL signs with mouthings (disambiguating mouth, where the mouthing distinguished between two manually identical signs—DM); and (4) BSL signs with echo phonology (EP).

The stimuli were designed to vary on the dimensions of presence or absence of mouth opening/ closing; presence or absence of hand and arm movements; and presence or absence of English-based mouth actions (**Table 4**).

Stimuli consisted of single words/signs, examples of which are given in **Table 5**. The list of silently spoken words was based on English translations of the signs.

Figure 3 shows examples of each of the stimulus types:

Thirteen (6 female; mean age 27.4; age range: 18–49) right handed participants participated. Volunteers were congenitally deaf native signers, having acquired BSL from their deaf parents. Stimuli were presented in alternating blocks of each of the experimental and a baseline condition. In order to encourage lexical processing, participants performed a target-detection task. Full

Table 4 | Characteristics of stimuli in fMRI experiment.

	Mouth opening and closing	Hand- arm movements (BSL)	English-derived mouth
No mouth (HO)	−	+	−
Phonology (EP)	+	+	−
Disambiguating mouth (DM)	+	+	+
Silent speech (SS)	+	−	+

Table 5 | Examples of stimuli in fMRI study (EP syllables in brackets).

EP	DM	Man
EXIST [ɯ]	FINLAND/METAL	TABLE
WIN [hy]	BATTERY/AUNT	CHERRY
NONE [pu]	WOOD/PROBLEM	BUTTER
SUCCESS [pa]	RUSSIA/BOY	KNOW
END [pəm]	ITALY/WIN	FAX

details of the experimental protocol and analysis may be found in Capek et al. (2008).

SIGN LANGUAGE (MAN, DM, EP)

In all three sign language conditions, Deaf native signers activated core language regions that are typically found when hearing people listen to speech. Although both sign language and speech involve perisylvian regions, sign language perception activated more posterior and inferior regions (**Figure 4**).
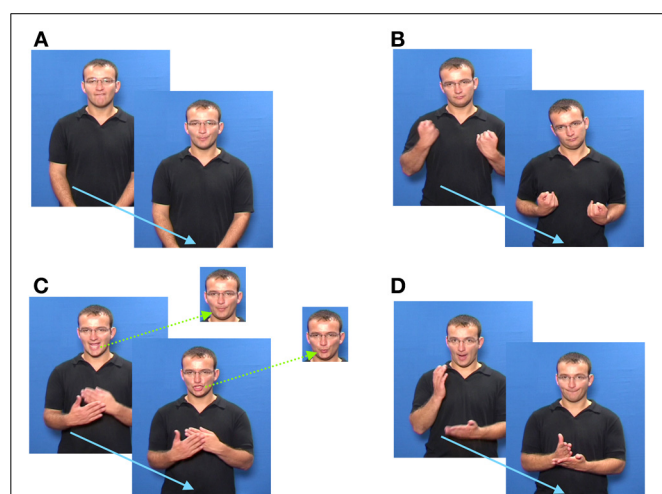

FIGURE 3 | Illustrations of stimuli. (A) SS, Silent articulation of the English word "football." The fricative (/f/)("foot.."), and the semi-open vowel /ɔː/ (".ball") are clearly visible. **(B)** Man, The BSL sign ILL. **(C)** DM, The BSL sign ASIAN shows the mouthing of /eɪ/ and /ʒ/. The face insets show the corresponding parts of the mouthings for the manual homonym BLUE, where /b/ and /uː/ can be seen. **(D)** EP, The manual sequence for [TRUE] requires abrupt movement from an open to a closed contact gesture. As this occurs, the mouth closes abruptly.

COMPARING ECHO PHONOLOGY (EP) AND OTHER MOUTHINGS (DM)

The task required participants to process material linguistically. In order to achieve lexical processing, BSL users must integrate perceptual processing of hands and of face/head, and this needs to be achieved fluently and automatically. If the cortical circuitry for sign language processing were driven by a mechanism that is "articulation-blind," we would expect there to be no systematic differential activation between signs with mouthings (where the mouth information is non-redundant), signs with no mouth action, and signs with echo phonology. Yet the contrasts found suggest this is not the case.

DM generated relatively greater activation in a circumscribed region of the left middle and posterior portions of the superior temporal gyrus (resembling the speech reading condition), while EP produced relatively greater posterior activation (Capek et al., 2008, p. 1231). We can consider the four conditions to represent a continuum from speech (SR) to speech accompanying signs (DM) to signs with accompanying non-speech-like mouth actions (EP) to purely manual signs (Man). Since greater posterior activation is characteristic of more sign-like material, EP also occupies an intermediate position between signs without mouth and signs with mouth actions derived from spoken language (**Figure 5**) in terms of neural activity.

The comparison of mouthings (DM) and echo phonology (EP) provides information about the nature of the mouth movements, and their role in sign language processing. The only differences in activation between DM and EP signs were found in the temporal lobe, with echo phonology (which is not derived from speech) demonstrating relatively greater posterior activation in both hemispheres than DM. This can be interpreted as a cortical correlate of the claim that the hands are indeed "the head of the mouth" (Boyes-Braem and Sutton-Spence, 2001), for echo phonology, as proposed by Woll (2001). While DM resembles speechreading in terms of functional cortical correlates, activation
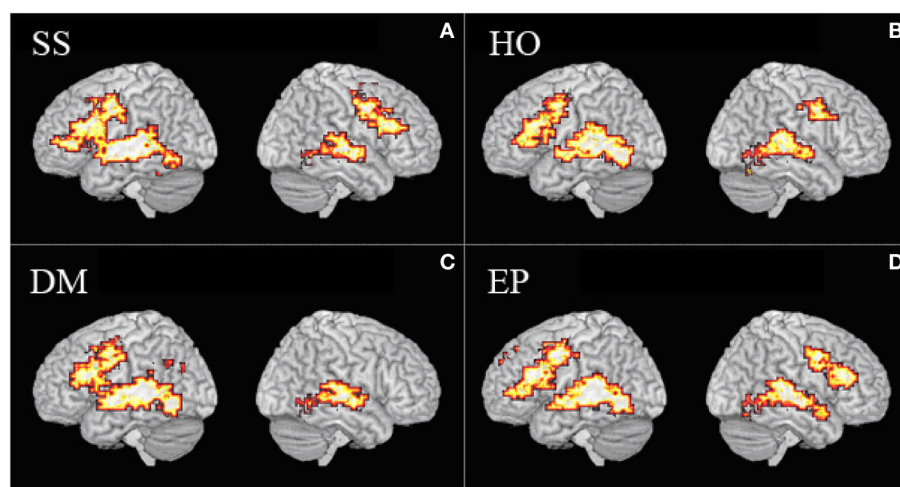

FIGURE 4 | Brain activation. (A) Activation during silent speechreading (SS). **(B)** Activation during processing of signs without any mouth actions (HO). **(C)** Activation during processing of signs with disambiguating mouth actions (DM). **(D)** Activation during processing of signs with echo phonology (EP).
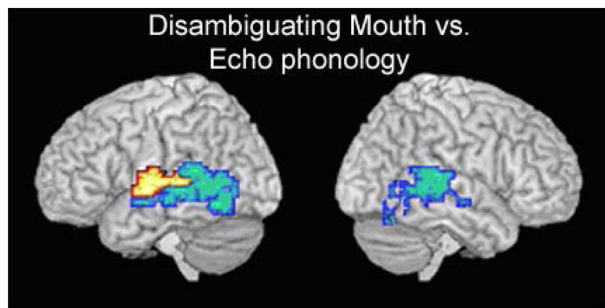
FIGURE 5 | Contrast between activation for DM (yellow) and EP signs (blue).

for EP resembles that for manual-only signs. Thus EP appears to occupy an intermediate position between spoken words and signs.

CONCLUSIONS

One issue for those concerned with suggesting a link between gesture and word has always been how the arbitrary symbol-referent relationship of words in spoken language could have come from visually-motivated gestures. Echo phonology provides evidence for a possible mechanism. Firstly, the phenomenon appears to be fairly common across different sign languages (although the occurrence of echo phonology remains to be researched in non-European sign languages). Secondly, the mouth actions found in echo phonology are themselves non-visually motivated. For example, signers report that BSL EXIST is iconic (Vinson et al., 2008), indicating "something located there," but it is impossible to reconstruct from the echo phonology syllable which accompanies it [ʃ] the meaning "exist," Thirdly, the actual inventory of elements in echo phonology looks very much like a system of maximal contrasts in a spoken language phonology (although there are some limitations because of the absence of sound contrasts). Fourthly, functional imaging research on the representation of signs and words in the brain suggests that echo phonology occupies an interesting intermediate position.

This paper represents a preliminary exploration of echo phonology. However, the data lead us to a number of conclusions. They support the arguments of those who argue against the notion that a unimodal manual protolanguage preceded the evolution of spoken language, since they demonstrate the extent to which signs are combined with mouth actions. The data also provide a window onto a mechanism by which the arbitrary pairing of a referent with a symbol (Saussure's defining feature of spoken language) could have occurred. Further research is needed to explore the presence of echo phonology in other sign languages (including those with a more recent point of creation than BSL) and whether echo phonology is subject to change (for example, added or transformed in a process of sign conventionalization). These studies may provide more insights into the origins of phonological/lexical structure in spoken language, and from that to the evolution of human language.

REFERENCES

Atkinson, J., Marshall, J., Woll, B., and Thacker, A. (2005). Testing comprehension abilities in users of British sign language following CVA. *Brain Lang.* 94, 233–248. doi: 10.1016/j.bandl.2004.12.008

Barbieri, F., Buonocore, A., Dalla Volta, R., and Gentilucci, M. (2009). How symbolic gestures and words interact with each other. *Brain Lang.* 110, 1–11. doi: 10.1016/j.bandl.2009.01.002

Bernardis, P., and Gentilucci, M. (2006). Speech and gesture share the same communication system. *Neuropsychologia* 44, 178–190. doi: 10.1016/j.neuropsychologia.2005.05.007

Boyes-Braem, P., and Sutton-Spence, R. L. (eds.). (2001). *The hands are the head of the mouth*. Hamburg: Signum Press.

Capek, C. M., Waters, D., Woll, B., MacSweeney, M., Brammer, M. J., McGuire, P. K., et al. (2008). Hand and mouth: cortical correlates of lexical processing in British Sign Language and speechreading English. *J. Cogn. Neurosci.* 20, 1220–1234. doi: 10.1162/jocn.2008.20084

Corballis, M. C. (2003). From mouth to hand: gesture, speech, and the evolution of right-handedness. *Behav. Brain Sci.* 26, 198–208. doi: 10.1017/S0140525X03000062

Corina, D. P., San Jose-Robertson, L., Guillemin, A., High, J., and Braun, A. R. (2003). Language lateralization in a bimanual language. *J. Cogn. Neurosci.* 15, 718–730. doi: 10.1162/jocn.2003.15.5.718

Crasborn, O. (2006). "Nonmanual structures in sign language," in *Encyclopedia of Language and Linguistics,* Vol. 8, ed K. Brown (Oxford: Elsevier), 668–672.

Crasborn, O., Mesch, J., Waters, D., Nonhebel, A., van der Kooij, E., Woll, B., et al. (2007). Sharing sign language data online: experiences from the ECHO project. *Int. J. Corpus Linguist.* 12, 537–564. doi: 10.1075/ijcl.12.4.06cra

Crasborn, O., van der Kooij, E., Waters, D., Woll, B., and Mesch, J. (2008). Frequency distribution and spreading behavior of different types of mouth actions in three sign languages. *Sign Lang. Linguist.* 11, 45–67. doi: 10.1075/ijcl.12.4.06cra

Darwin, C. (1872). *The Expression of the Emotions in Man and Animals.* London: John Murray. doi: 10.1037/10001-000

Emmorey, K. (2001). *Language, Cognition and the Brain.* Hove: Psychology Press.

Erhard, P., Kato, T., Strick, P. L., and Ugurbil, K. (1996). Functional MRI activation pattern of motor and language tasks in Broca's area. *Soc. Neurosci.* Abstracts 22:260.2.

Fitch, W. T. (2010). *The Evolution of Language.* Cambridge: Cambridge University Press. doi: 10.1017/CBO9780511817779

Gentilucci, M. (2003). Grasp observation influences speech production. *Eur. J. Neurosci.* 17, 179–184. doi: 10.1046/j.1460-9568.2003.02438.x

Gentilucci, M. C., and Corballis, M. (2006). From manual gesture to speech: a gradual transition. *Neurosci. Biobehav. Rev.* 30, 949–960. doi: 10.1016/j.neubiorev.2006.02.004

Goldin-Meadow, S. (2003). *The Resilience of Language: What Gesture Creation in Deaf Children can Tell us About how all Children Learn Language.* New York, NY: Psychology Press.

Hewes, G. W. (1973). Primate communication and the gestural origin of language. *Curr. Anthropol.* 14, 5–24. doi: 10.1086/201401

Kendon, A. (2010). *A Discussion of the Evolution of Language by Tecumseh Fitch. Semiotix XN-3.* Available online at: http://www.semioticon.com/semiotix/files/2010/11/Fitch-Essay.pdf

Klima, E. S., and Bellugi, U. (1979). *The Signs of Language.* Cambridge, MA: Harvard University Press.

Liberman, A. M., and Mattingly, I. G. (1985). The motor theory of speech perception revised. *Cognition* 21, 1–36. doi: 10.1016/0010-0277(85)90021-6

Liddell, S. K. (1978). "Nonmanual signs and relative clauses in American Sign Language," in *Understanding Language through Sign Language Research,* ed P. Siple (New York, NY: Academic Press), 59–90.

MacNeilage, P. F. (1998). The frame/content theory of evolution of speech production. *Behav. Brain Sci.* 21, 499–546. doi: 10.1017/S0140525X98001265

MacSweeney, M., Capek, C., Campbell, R., and Woll, B. (2008). The signing brain: the neurobiology of sign language. *Trends Cogn. Sci.* 12, 432–440. doi: 10.1016/j.tics.2008.07.010

MacSweeney, M., Woll, B., Campbell, C., McGuire, P. K., Calvert, G. A., David, A. S., et al. (2002). Neural systems underlying British Sign Language sentence comprehension. *Brain* 125, 1583–1593. doi: 10.1093/brain/awf153

Mather, S., and Malkowski, S. (2013). "Examining echo phonology in American Sign Language (ASL)," in *Poster Presented at Theoretical Issues in Sign Language Research XI* (London).

McNeill, D., Duncan, S., Cole, J., Gallagher, S., and Bertenthal, B. (2008). "Growth points from the very beginning," in *Interaction Studies: Special Issue on Protolanguage 9(1),* eds M. Arbib and D. Bickerton (Amsterdam: John Benjamins), 117–132.

Meguerditchian, A., and Vauclair, J. (2008). "Vocal and gestural communication in nonhuman primates and the question of the origin of language," in *Learning From Animals? Examining the Nature of Human Uniqueness,* eds L. Roska-Hardy and E. M. Neumann-Held (London: Psychology Press), 61–85.

Newman, A. J., Supalla, T., Hauser, P. C., Newport, E. L., and Bavelier, D. (2010). Prosodic and narrative processing in American Sign Language: an fMRI study. *Neuroimage* 52, 669–676. doi: 10.1016/j.neuroimage.2010.03.055

Paget, R. (1930). *Human Speech—Some Observations, Experiments and Conclusions as to the Nature, Origin, Purpose and Possible Improvement of Human Speech.* London: Kegan Paul, Trench, Trubner & Co.

Pendzich, N.-K. (2013). "Lexical layering of manual and nonmanual components within specific signs: an empirical study of German Sign Language (DGS)," in *Poster Presented at Theoretical Issues in Sign Language Research XI* (London).

Rizzolatti, G., and Arbib, M. A. (1998). Language in our grasp. *Trends Neurosci.* 21, 188–194. doi: 10.1016/S0166-2236(98)01260-0

Rizzolatti, G., and Craighero, L. (2004). The mirror-neuron system. *Annu. Rev. Neurosci.* 27, 169–192. doi: 10.1146/annurev.neuro.27.070203.144230

Rizzolatti, G., Fadiga, L., Gallese, V., and Fogassi, L. (1996). Premotor cortex and the recognition of motor actions. *Cogn. Brain Res.* 3, 131–141. doi: 10.1016/0926-6410(95)00038-0

Stokoe, W. C. (1960). *Sign Language Structure.* Buffalo, NY: University of Buffalo.

Sutton-Spence, R., and Woll, B. (1999). *The Linguistics of BSL: an Introduction.* Cambridge: Cambridge University Press.

Swadesh, M. (1971). *The Origin and Diversification of Language,* reprinted 2006. Chicago, IL: Alldine Atherton.

Sweet, H. (1888). "The history of English sounds," in *The Indispensable Foundation: a Selection from the Writings of Henry Sweet,* ed E. J. A. Henderson (Oxford: Oxford University Press).

Vinson, D. P., Cormier, K., Denmark, T., Schembri, A., and Vigliocco, G. (2008). The British Sign Language (BSL) norms for age of acquisition, familiarity and iconicity. *Behav. Res. Methods* 40, 1079–1087. doi: 10.3758/BRM.40.4.1079

Woll, B. (2001). "The sign that dares to speak its name: echo phonology in British Sign Language (BSL)," in *The Hands are the Head of the Mouth,* eds P. Boyes-Braem and R. L. Sutton-Spence (Hamburg: Signum Press), 87–98.

Woll, B. (2009). "Do mouths sign? Do hands speak?: Echo phonology as a window on language genesis," in *Language Evolution: the View from Restricted Linguistic Systems,* eds R. Botha and H. Swart (Utrecht: LOT Occasional Series), 203–224.

Woll, B. (2012). "Atypical signing," in *Sign Language: an International Handbook,* eds R. Pfau, M. Steinbach, and B. Woll (Berlin: Mouton de Gruyter), 762–787.

Woll, B., and Sieratzki, J. S. (1998). Echo phonology: signs of a link between gesture and speech. *Behav. Brain Sci.* 21, 531–532. doi: 10.1017/S0140525X98481263