Temporal Attention as a Scaffold for Language Development

Ruth de Diego-Balaguer[1,2,3]*, Anna Martinez-Alvarez[2,3] and Ferran Pons[3,4]

[1] Institució Catalana de Recerca i Estudis Avançats, Barcelona, Spain, [2] Cognition and Brain Plasticity Unit, Institut d'Investigació Biomèdica de Bellvitge, Barcelona, Spain, [3] Department of Basic Psychology, University of Barcelona, Barcelona, Spain, [4] Department of Basic Psychology, Institute for Brain, Cognition and Behavior (IR3C), University of Barcelona, Barcelona, Spain

Language is one of the most fascinating abilities that humans possess. Infants demonstrate an amazing repertoire of linguistic abilities from very early on and reach an adult-like form incredibly fast. However, language is not acquired all at once but in an incremental fashion. In this article we propose that the attentional system may be one of the sources for this developmental trajectory in language acquisition. At birth, infants are endowed with an attentional system fully driven by salient stimuli in their environment, such as prosodic information (e.g., rhythm or pitch). Early stages of language acquisition could benefit from this readily available, stimulus-driven attention to simplify the complex speech input and allow word segmentation. At later stages of development, infants are progressively able to selectively attend to specific elements while disregarding others. This attentional ability could allow them to learn distant non-adjacent rules needed for morphosyntactic acquisition. Because non-adjacent dependencies occur at distant moments in time, learning these dependencies may require correctly orienting attention in the temporal domain. Here, we gather evidence uncovering the intimate relationship between the development of attention and language. We aim to provide a novel approach to human development, bridging together temporal attention and language acquisition.

Keywords: language development, infancy, attention, temporal orienting, statistical learning, rule learning, morphosyntactic development, word segmentation

INTRODUCTION

Speech is a complex auditory stimulation. A single word in speech can be perceived as a sequence of phonemes, as a whole word, as a stem and a suffix or as having a specific meaning, depending on the level of processing. In order to face this complexity infants do not learn all of this information at once but rather in an incremental fashion. In particular, two main linguistic milestones -word segmentation and non-adjacent rule acquisition- appear in a sequential fashion. During the first months, infants are able to segment speech into words and recognize them; however, it is not until after the first year that they are able to understand and detect the subtle changes carried by different rule transformations (Gómez and Maye, 2005; Christophe et al., 2008). As a matter of fact, brain development in general is not uniformly distributed through infancy. Different brain structures and their white matter connections do not develop homogenously, nor do all cognitive functions develop at the same speed (Gogtay et al., 2004; Diamond, 2007). In particular attention shows also a developmental progression with different mechanisms arising at different moments

of development. *Exogenous attention*, captured by salient events in the environment, is functional much earlier than *endogenous attention*, which allows for selecting which information to process and which to ignore (Posner and Cohen, 1984). The progressive general cognitive development may be seen as affecting all functions independently. However, cognitive functions do not work in isolation. In particular, the attentional system acts as a filter to any incoming stimulation, influencing perception, and therefore may affect learning, which suggests that the development of the attention system is likely to shape the way language is processed and how it develops along with the available attention resources.

Extensive literature has previously reported a link between attention and different aspects of language processing. Adult speakers of different languages tend to adapt the syntactic structure of their productions as a function of their focus of attention on the visual information they describe (Myachykov et al., 2005; Ibbotson et al., 2013; Tomblin and Myachykov, 2015). Similarly, focus and topicalization are naturally used to draw attention to relevant elements in the sentence (Jackendoff, 2002). Indeed, using focus to comprehend sentences activates left inferior parietal lobe (IPL) overlapping with the attention-orienting network used in visual attention (Kristensen et al., 2012). Even babies produce isolated words in an attempt to attract the listener's attention to their own focus of interest (Jackendoff, 2002). This behavior is closely linked to the development of joint attention present at the end of the first year of life (Bruner, 1983; Carpenter et al., 1998; Brooks and Meltzoff, 2002). From joint attention interaction infants begin to link words with objects and events (Baldwin, 1991, 1993).

This previous literature converges with our proposal with the close link between attention and language, the source of the link is closely related in these studies to the communicative roots of language, given that attention is used as a tool to drive the listener's attention to the focus of attention of the speaker (Smith et al., 2011). However, our proposal, in contrast, is intimately related to processing. The development of attention affects how the input is processed because it filters the input received, independently of the presence of an interlocutor and in the absence of a message to be transmitted. This is what shapes language learning from this point of view, in the same way as it shapes learning of other sources of information.

The present proposal presents an integrative approach of language acquisition, in which the powerful and dynamic interplay of exogenous and endogenous attention mechanisms allows infants to focus on different aspects of speech at different moments in development. In particular, during the initial stages of language acquisition, attention is captured by salient elements of speech, such as prosodic cues (e.g., pitch, rhythm, or pauses) because the infant perceptual system is guided by stimulus-driven attention. As months pass and endogenous attention progressively develops, this more flexible mechanism can be used to learn non-adjacent linguistic dependencies. This allows filtering out irrelevant information and selectively focusing on relevant elements that reliably predict forthcoming information. As we will argue and support with evidence from infant

development, changes in these two different aspects of cognition are not independent. In other words, the attention mechanisms available early on limit the type of linguistic information that infants can extract from speech. The delay in development of more controlled mechanisms of endogenous attention may not indicate a disadvantage in language acquisition but rather an advantage at early linguistic stages. That is, in agreement with Newport's "Less is More" hypothesis (Newport, 1990), this delay in the development of endogenous attention and the initial use of more automatic exogenous attention mechanisms may allow young infants to face a perceptual simplification of the complex speech stream early in the learning process. Therefore, relying in prosodic cues may be crucial and beneficial during the first months of life, when exogenous attention is the main mechanism available. Such a pattern would lead to the observed early segmentation and acquisition of words during the first months, which is followed by a shift later in infancy to focus on the upcoming information indicated by relevant cues (as in non-adjacent dependency learning) when the infant is able to select which information to attend and which to ignore.

Crucially, the other important difference from the previous theoretical approaches is that our proposal is based on the allocation of attention in the *temporal domain*. Whereas previous proposals have focused purely on visual attention and how it influences the conceptualization of the message to be conveyed (Levelt, 1989) and how this is reflected in the linguistic output, we are interested in how attention affects the processing of the ongoing auditory stimulation -the speech flow. Given that speech is a sequence of sounds that unfolds in time, attention to speech is necessarily oriented in the temporal domain. Because temporal-selective attention directs resources to certain moments in time, enhancing perception (Correa et al., 2006; Nobre et al., 2007), it can allow for the extraction of different events in speech (e.g., consonants, vowels, words, and phrases) that have different durations and appear in a certain order and moments in ongoing speech. Recently, more general proposals have also underscored that cerebral mechanisms for timing and ordinal knowledge are in charge of the neural representation of sequences in different domains, including language (Dehaene et al., 2015). More precisely, speech has temporally rhythmic and salient prosodic cues that capture attention automatically when they appear, helping, for example, to locate boundaries to segment words. On the other hand, segments carrying cues for rule dependencies may have different durations and different onset times (such as suffixes and pronouns). Attention can be progressively tuned to focus on these cues when they are progressively noticed to predict later upcoming dependencies in a sequence of words. This tuning requires the engagement of endogenous attention in the temporal domain. Therefore, acquiring words and rules may require the engagement of different attentional systems. A dynamic shift between systems should develop in the course of learning. Indeed, recent data in adults show that the same prosodic cues can lead to exogenous effects related to segmentation, even in the absence of any possible learning and endogenous effects when the prosodic information can be used as a cue to extract non-adjacent rules (de Diego-Balaguer et al., 2015).

By studying the attention mechanisms involved in language learning, we also pave the way to understand some of the sources of language learning disabilities. In the following sections, we first describe the typically developmental trajectory of attention and language functions and their underlying brain development. We review evidence supporting the hypothesis that maturation of the attention mechanisms may serve as a scaffold for language development, and we review evidence indicating the close relationship between attention deficits and impairments in language acquisition.

STAGES OF DEVELOPMENT OF THE ATTENTION SYSTEM

Before entering into the details of attention development, some conceptual clarifications can be helpful concerning the terms that are used throughout this paper. In the attention literature, a distinction between exogenous (bottom–up) and endogenous (top–down) attention has been classically proposed, and a plethora of studies have dissected the effects that characterize each of these systems and their interactions (Chica et al., 2013). In brief, both types of attention have been proven to facilitate processing. However, exogenous orienting appears even when a secondary task is performed, and it can be voluntary attenuated but not completely suppressed. Endogenous attention is often voluntary, but it can also appear with no effort and even when participants are not aware of the relationship between the cue and the target. Some models of attention (Corbetta and Shulman, 2002; Corbetta et al., 2008) propose a distinction in terms of *stimulus-driven* vs. *goal-directed* attention, which partially overlap with *exogenous* vs. *endogenous* attention but have important discrepancies that are worth mentioning here. Within the frontal and parietal brain regions involved in attention, stimulus-driven attention involves a more ventral fronto-parietal network (Corbetta et al., 2008), including the inferior parietal cortex, the ventral and inferior prefrontal cortex (PFC), and insula, as well as subcortically, the superior colliculus. Goal-directed attention, in contrast, involves a more dorsal fronto-parietal network, including the middle PFC and the superior parietal lobe (Corbetta et al., 2008), and subcortically, the pulvinar of the thalamus.

An important distinction within this framework that might help to understand the attentional systems in a less dichotomic way is the distinction between saliency when (i) no task or goal is present (i.e., exogenous saliency) compared to when (ii) elements are salient because they share some feature that is relevant for the task or goal of the subject despite not being the target of the task (i.e., task-relevant saliency). In other words, a red circle surrounded by green squares will attract attention in the absence of any task due to their exogenous saliency; however, a green circle will drive our attention if our task is to detect a green square because the circle shares a relevant feature (i.e., green) to our task. This distinction is important not only because neuroimaging data show that when performing a task, sensory-salient and task-relevant stimuli induce the activation of different brain networks but also because in the absence

of a task, these different types of stimulation do engage the ventral network (Corbetta and Shulman, 2002; Chica et al., 2013). Because very young infants do not have a goal-directed system available, salient stimuli in the environment may trigger the ventral network and subcortical areas. With incremental learning and the progressive availability of the goal-directed system, the relevant elements in the environment attract the ventral attention system in a more *task-relevant* manner. In terms of what this might mean for language, infants may first be attracted by any change in pitch, pauses or in voice onset time in speech sounds due to their intrinsic saliency, whereas later in development, once prosodic characteristics, representations of the phonemes and words of their native language are learned, only those prosodic variations and speech sounds that correspond to the contrast of their language will attract their attention. To avoid misunderstandings in the course of the paper, we will refer to *saliency* to designate only those stimuli that attract attention due to their sensory characteristics irrespective of their relevance.

Turning back to the development of the attentional system, three main attentional mechanisms have been described in the literature on attention development: alertness, orienting, and endogenous attention (Colombo, 2001). The first two characterize exogenous attention. Rudimentary forms of each of the functions of attention are already present to some degree at birth, but each exhibits progressive maturity during the first years of life.

The arousal system is already present at early stages of development. This attention system is associated with an infant's level of alertness and readiness to process stimuli from the environment. From birth to 2 months of age, alertness is commonly initiated by exogenous stimulation (Wolff, 1965). Very young infants show "obligatory attention" (Stechler and Latz, 1966) or "sticky fixation" (Hood, 1995), a difficulty in interrupting gaze from a given stimulus they are fixating in order to shift attention to a different one. The efficiency of disengaging from and shifting gaze to a stimulus increases during the first months after birth (Hunnius and Geuze, 2004). This phenomenon is tied to the neurological maturation of the visual pathway and associated with subcortical structures (Richards et al., 2010) that, although present at birth, are still developing in terms of their connectivity to most cortical areas (Casey et al., 2004; Uylings, 2006). Between 2 and 3 months of age, maturation moderates the inhibition mechanisms that limit eye-movements, which start to gain cortical control. Effective visual exploration requires disengaging and shifting gaze across different locations, and the perseveration of this sticky phenomenon at 7 months of age is an early feature of later emerging autism (Elison et al., 2013) that persists with disengagement difficulties in childhood (Landry and Bryson, 2004). Therefore, this attentional mechanism might be more related to communicative and social aspects of language development.

At later stages, infants start developing the ability to orient attention toward a particular stimulus in space (Courage and Richards, 2008). The infant's visual behavior in the first year of life is dominated by an orienting system of attention with two main components (Ruff and Rothbart, 1996). On the one hand, the

spatial-orienting network, which includes the posterior parietal cortex and several subcortical systems, mediates attentional functions, such as engagement, disengagement, shifting, and inhibition of return. On the other hand, the object recognition network, which includes pathways from the primary visual cortex to the parietal cortex and the inferior temporal cortex, orients attention to object features. A remarkable developmental progression of this orienting system occurs between 3 and 9 months of age (Ruff and Rothbart, 1996). During this period infant flexibly and quickly orient attention to stimuli in the environment in terms of experiential factors (e.g., novelty or complexity) rather than their exogenous salience (Courage et al., 2006).

Finally, the endogenous orienting of attention shows a slower and later developmental time course than other attention systems, showing a remarkable change during the later parts of the first year and beyond (Colombo, 2001). It is not until the end of the first year that more complex features arise and endogenous control of attention starts acquiring an executive component to a greater extent (Courage and Richards, 2008), which is closely related to the initial maturation of the dorsal prefrontal and the anterior cingulate cortices (Posner and Petersen, 1990). Age-related attentional improvements are related to changes in structural and functional connectivity (Rueda et al., 2015). Data reveal that increased attentional performance is related to greater information transfer in the brain, which involves distributed brain nodes and paths that connect these nodes. Crucially, the anterior cingulate cortex does not begin to develop long-range connectivity with other brain areas until after the first year of life, developing progressively during childhood (Fair et al., 2009; Gao et al., 2009). However, although the structural connectivity pattern in children resembles that of adults, the functional connectivity of attentional networks shows different patterns. While adults' orienting and executive attentional systems exhibit separate functional networks, these systems are more unified in children (Fair et al., 2007). In summary, whereas exogenous attention shows an earlier maturation course, endogenous attention develops later and slower, continuing its development through childhood and until adolescence (Colombo and Cheatham, 2006; Johnson et al., 2015).
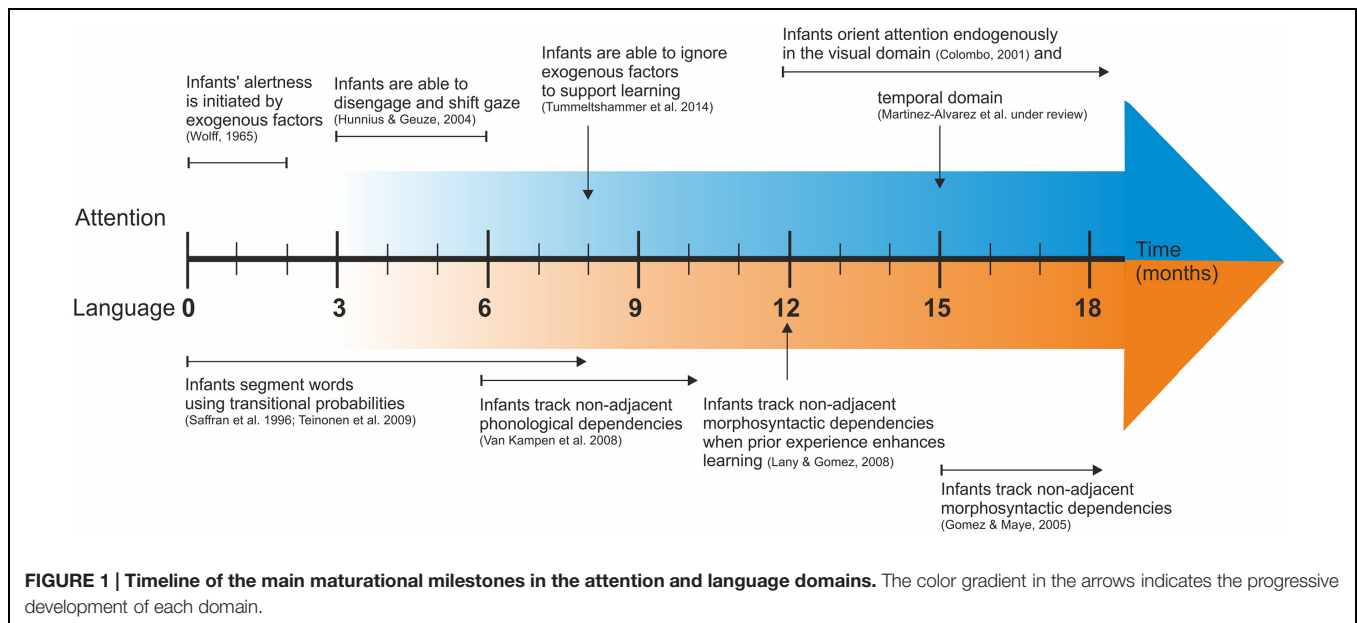
Despite the apparent dichotomy between exogenous and endogenous attention, the appearance of the latter does not imply an inhibition of the exogenous mechanism but a better interaction between both. As it has been highlighted in the adult literature (Corbetta et al., 2008), stimulus-driven attention is able to break the engagement of goal-directed attention, highlighting the close interaction between systems. However, goal-directed attention can attenuate the interference from distractors by decreasing the activation of stimulus-driven attention. The ability to ignore salient distractors to support learning is observed at 8 and 12 months (Althaus and Mareschal, 2012; Tummeltshammer et al., 2014) but is not present earlier, indicating a stronger influence of exogenous factors on young infants' attention. From this point of view, it is clear that the development of the exogenous and endogenous mechanisms of attention do not show a strict sequential order but rather a smooth overlap with subtle signs of endogenous attention appearing before 8 months of age but with poor command observed to progressive control reached at the end of the first year (see **Figure 1**).

Most of these developmental descriptions of infants have been based on studies of visuospatial attention. Although this description is useful in understanding the developmental progression of attention in general, tracking auditory information in time is critical when we consider language learning due to the intrinsic temporal characteristics of speech. Generally speaking, to orient attention in time, different capacities need to be in place. Infants need to be able to perceive the difference between distinct temporal lags and be sensitive to the order of elements in a sequence. Once these perceptual capacities are available, the attentional system should be able to orient to these different elements in the auditory domain to extract information from speech. Temporal processing is supported by a cortico-subcortical network, including the premotor cortex, basal ganglia and cerebellum that have been proposed to also be involved in speech processing (Kotz and Schwartze, 2010). As previously mentioned, subcortical structures are functional and are used by infants from birth (Casey et al., 2004; Uylings, 2006), and sensorimotor cortices are the first to develop (Dehaene-Lambertz and Spelke, 2015). The early availability of these structures may allow infants to use temporal information in an exogenous manner in the early stages of development.

Studies exploring infants' ability to perceive time mostly focus on regular temporal structure perception (e.g., rhythm and regular isochronic sequences), i.e., focusing in attention mechanisms that are mainly exogenous and stimulus-driven (Demany et al., 1977; Haith et al., 1988; Adler et al., 2008). These studies show that in the first months of age, infants are sensitive and can orient their attention in time following regular patterns. vanMarle and Wynn (2006) reported that 6-month-old infants can discriminate event durations between 2 and 4 s. Brannon et al. (2004, 2008) additionally showed that 10-month-old infants can detect changes in temporal rhythm by detecting a temporal deviation in a stream of tones formed by a regular inter-stimulus interval. In terms of infants' ability to benefit from rhythmic and regular patterns, their behavior is similar to that observed in adult research (Large and Jones, 1999; Barnes and Jones, 2000; Sanabria et al., 2011).

In contrast, the ability to orient attention in time endogenously has not been reported in infancy. Recent data (Martinez-Alvarez et al., under review) indicate that the ability to endogenously orient attention in time appears after the ability to orient attention in space. More precisely, whereas 12-month-olds show only spatial orienting abilities, 15-month-olds are able to adapt their anticipatory behavior according to both spatial and temporal predictive cues. A recent study with children revealed also that the developmental trajectory of voluntarily use temporal cues is delayed relative to the use of spatial cues. However, this study showed that 11-year-olds were only able to implicitly but not voluntarily orient attention in time (Johnson et al., 2015), which apparently seems to contradict the results with infants (Martinez-Alvarez et al., under review). Different explanations of these results with children are possible. As Johnson and colleagues explain, one possibility is that the temporal cues were

FIGURE 1 | Timeline of the main maturational milestones in the attention and language domains. The color gradient in the arrows indicates the progressive development of each domain.

conceptually more demanding than the spatial cues. Another possibility is that the spatial uncertainty of target appearance in their paradigm diminished the utility of the temporal cue. Electrophysiological and behavioral investigations have shown that temporal predictability is most successful when joined with spatial predictability (Doherty et al., 2005; Rohenkohl et al., 2014). Indeed, preliminary evidence from Coull's lab shows that children can use temporal cues when the spatial location of the target is known in advance (Johnson et al., 2015). Further investigations are needed for a better understanding the development of temporal attention at the functional and anatomical levels.

In sum, the development of attention is characterized by a shift from exogenous, stimulus-driven orienting of attention, particularly during the first 3 months of age, to a smooth progression to greater endogenous control, the first hints observable before 8 months and showing a marked dominance after the first year (Johnson, 1990; Ruff and Rothbart, 1996). Although little evidence is available from attention orienting in the temporal domain, clear effects of sensitivity to temporal differences and rhythmic cues are present early on, whereas endogenous orienting of attention in time appears later.

LANGUAGE DEVELOPMENT IN INFANCY IN RELATION TO ATTENTION

In the current section, we review the developmental trajectory of linguistic abilities in infants, focusing on studies on word segmentation and non-adjacent rule learning, which are the two main milestones of interest for our hypothesis. Throughout the review, we point out the role of the attention mechanisms related to the language data available at each stage.

Early Stages of Language Learning and Exogenous Mechanisms of Attention

The early capacities of infants to acquire their native language have been extensively reported. Even before they begin to produce their first words, infants have already acquired an important amount of linguistic knowledge. One very early ability is their sensitivity to perceive the rhythmic characteristics of language at birth, showing discrimination of the stress patterns in different languages and an early preference for their native language stress pattern (Nazzi et al., 1998). Importantly, prosodic characteristics, such as intonation, stress and pitch variations, are salient perceptual cues that can easily attract infant's exogenous attention. These prosodic cues play a key role in word segmentation because infants exploit these even before they use other cues to locate word boundaries (Mattys et al., 1999).

During the first months of life, infants are capable of extracting words in spoken language by detecting and exploiting other perceptual cues. For example, neonates (Teinonen et al., 2009) and 8 months old infants can make use of statistical regularities between adjacent syllables (also known as transitional probabilities, TP) to locate word boundaries and extract words from both artificial (Saffran et al., 1996) and natural languages (Pelucchi et al., 2009). On the other hand, the combination of both prosodic and statistical cues shows how predominant are the acoustic features of infant-directed (ID) speech (e.g., exaggerated pitch contours) compared to adult-directed (AD) speech to attract infants' attention (Fernald, 1985; Cooper and Aslin, 1990) and to facilitate infants' word segmentation (Trainor and Desjardins, 2002; Thiessen et al., 2005).

Statistical learning is a remarkable ability, and numerous studies have been developed to understand the mechanisms underlying and the factors affecting this type of learning. Indeed, several important features of growing literature with the same paradigms are important to mention in relation to the hypothesis

outlined here. One critical factor is that statistical learning is a simple adaptive capacity that can also be found in other animals that have much less developed prefrontal cortices. For example, rodents exposed to the same type of linguistic speech streams are able to correctly segment it, albeit with a somewhat different computation (Toro and Trobalón, 2005). Another important feature is that the presence of statistical regularities in the input captures attention (Turk-Browne et al., 2005). Thus, even when no effort to learn is given, regularities can be extracted from the input (Saffran et al., 1997), capturing our attention in an automatic manner, which is consistent with the fact that even newborns are able to detect these statistical regularities (Teinonen et al., 2009). Therefore, the development of endogenous attention is not necessary for this learning to occur. In the same vein, electrophysiological evidence indicates that once words are segmented, the recognition/detection of a known word within the speech stream also captures attention (Sanders et al., 2002; Parise et al., 2010; de Diego-Balaguer et al., 2015), enhancing the long-term memorization of the segmented word forms.

Another important fact underscores the importance of exogenous attention in these early learning stages and highlights the adaptive function of the unavailability of the endogenous system in young infants. In adults, the manipulation of diverted attention, orienting endogenous attention outside the speech stream, can interfere with teach (Toro et al., 2005). Adults and older infants can orient their attention endogenously, and although this can be helpful when it converges to track the critical information for learning, it can also interfere with learning when it diverts from the correct focus of attention. For example, attention diverted from the dependency by the attraction of novel words that need to be ignored prevents non-adjacent learning. Infants' ability to generalize the detection of non-adjacent dependency to nonsense stems (e.g., *These meeps*) occurs only if they are first presented with familiar stems (e.g., *These chairs*). When attention is captured by a *novel* intervening element, learning of non-adjacent dependencies is altered (Soderstrom et al., 2002). In contrast, if salient information automatically captures infants' attention and this information is helpful for learning, the absence of endogenous attention prevents infants from disengaging and reorienting their attention to a different focus of attention that may interfere with the correct computation. In this way, the early dominance of this automatic exogenous mechanism can make learning more likely to occur. Other salient features, such as adjacent repetitions, can also act as important attentional attractors improving learning. Already present at birth, infants possess an automatic perceptual mechanism to detect repetitions in the auditory domain (Endress et al., 2009). This is reflected in greater activation in the temporal and left frontal brain areas when tested for recognition after exposure to simple repetition-based structures (ABB; e.g., "mubaba," "penana") than to random sequences (ABC; e.g., "mubage," "penaku").

Overall, the evidence indicates that the characteristics of speech with their statistical regularities and the salient prosodic cues are perfectly adapted to make the most of the early availability of exogenous attention. By engaging exogenous stimulus-driven attention, available since birth, learning can be achieved. The absence of control of voluntary attention at these early stages of development does not limit infants' ability to acquire language but rather helps them by allowing infants to follow their stimulus-driven mechanism to capture the relevant information for learning automatically.

Early Signs of Non-Adjacent Dependency Learning in the Rise of Endogenous Attention

Although the ability to segment and extract words from speech is a critical milestone of language acquisition, to acquire the grammar of their language infants must also track non-adjacent relationships. Importantly, the extraction of hierarchical structures relies on temporally distant relationships and is fundamental to capture the properties of language (Chomsky, 1957). Nonadjacent dependencies refer to cases in which two elements co-occur over one or more intervening elements. In natural languages, for example, in English, there is an association between auxiliaries and inflectional morphemes, irrespective of the intervening verb stem (e.g., **is** walk**ing**; **is** runn**ing**; **is** eat**ing**). Infants must dismiss the variable irrelevant information and focus instead on the invariant relevant cues that predict the non-adjacent dependency (Gómez and Maye, 2005).

Because the endogenous system appears progressively in the course of development, its initial use in its earliest stages depends on the convergent presence of exogenous cues. The first signs of non-adjacent tracking in language are observed in the phonological domain where the presence of exogenous cues helps infants to track the dependencies grouped by their high similarity (for a review, see Sandoval and Gómez, 2013). For example, infants as young as 7 months can use harmony on vowel which are more salient than consonants and linked to prosodic variations as a cue to find word boundaries (Mintz and Walker, 2006; Kanpem et al., 2008) but cannot use consonantal harmony (Nazzi et al., 2009; Gonzalez-Gomez and Nazzi, 2012). They need to reach 10 months of age before they can, an age where endogenous attention starts to be more prominent. In a similar vein, newborns can discriminate adjacent rules based on the repetition of the same syllable but not when rules are non-adjacent (ABA; e.g., "bamuba," "napena") (Gervain et al., 2008). In contrast, 7-month-old and older infants track non-adjacent dependencies but only under some circumstances; when non-adjacent syllables are identical and the interleaved syllables are different (e.g., *le di le, ga po ga)* (Marcus et al., 1999; Gerken, 2006). Unexpectedly, a more recent study demonstrated that German infants as young as 4 months of age could discriminate between grammatical and ungrammatical non-adjacent dependencies in Italian (Friederici et al., 2011). As the authors indicate, Italian morphosyntactic dependencies also contain phonological dependencies. Given that phonological, non-adjacent dependencies are tracked from very early stages in development, it has been proposed that 4-month-olds may be tracking the phonological aspects to discriminate these non-adjacent dependencies. In other words, the exogenous attentional resources already available to 4-month-olds could have driven the success of such young infants in this task.

Learning more Challenging Non-Adjacent Dependencies with Greater Maturation of Endogenous Attention

Although young infants can track non-adjacent linguistic rules under certain learning conditions (e.g., when the dependent units are similar and the intermediate elements are dissimilar), learning of morphosyntactic dependency appears several months after phonological dependency learning has occurred. This non-adjacent dependency learning that appears in more challenging perceptual and linguistic arrangements, requires greater involvement of endogenous mechanisms. Simply tracking non-adjacent dependencies can be used to locate word boundaries but is not enough to extract and generalize the underlying rule that entails the creation of abstract categories for generalization (Peña et al., 2002).

In this context, prosodic information in natural languages provides reliable cues not only for word segmentation but also for rule learning (Jusczyk, 2002) because prosodic pauses tend to co-occur with syntactic boundaries. Nevertheless, although prosodic cues play a role in word segmentation from birth, it is not until the first year that infants start exploiting these cues for rule extraction (Johnson, 2008; Seidl and Johnson, 2008). The presence of this prosodic information in an artificial language enhances the extraction of non-adjacent dependencies compared to continuous speech streams without pauses (Peña et al., 2002). An important point to highlight is that the presence of pauses *per se* does not improve learning of non-adjacent dependencies. Those pauses need to occur at the boundaries of the position of the dependencies to be useful (Endress et al., 2005; Mueller et al., 2010). The use of these cues (stress pattern or prosodic pauses) for word segmentation requires only orienting attention to the position of the prosodic information that captured attention, that is, in an exogenous fashion. However, the use of prosodic pauses for rule extraction additionally requires the use of this cue to selectively focus attention on concurrent phonological information at this specific position. This cue then has to be used as a relevant predictor of forthcoming information to extract the rule dependency, which implies focusing attention to this cue and the predicted element while disregarding the intervening irrelevant information.

Within the morphosyntactic domain, nonadjacent relationships are often found between subject and verb agreement (**he** walk**s**) or between auxiliary and verb agreement (he **is** walk**ing**). Learning of morphosyntactic, non-adjacent dependencies emerges after the first year of life (Gómez and Maye, 2005). This developmental course is reasonable when considering the challenge of the task, that is, in order to track the dependency among non-adjacent elements, infants must first identify the morphemes without involving any given similarity and then track the dependency between them across intervening elements irrelevant to the rule dependency (**he** walk**s**; **he** run**s**; **he** eat**s**). In one of the first studies exploring infants' ability to learn verb–tense agreement (Santelmann and Jusczyk, 1998), researchers reported that 18-month-olds accepted grammatical phrases in English, such as "is running,"

and rejected ungrammatical phrases, such as "can running," whereas 15-month-olds were not able to differentiate between the phrases. Moreover, learning was possible only under certain conditions, with infants succeeding when the intervening element extended three syllables or less (e.g., *Grandma **is** always sing**ing***, but not *Grandma **is** almost always sing**ing***).

In addition, in order to learn a non-adjacent relation of the form "**these** cats**,**" infants must track a dependency between two elements that occur over an intervening element and create different categories (e.g., determiner, noun, verb). Several lines of research have explored the mechanisms underlying the ability of grouping elements into categories. For example, it has been proposed that *frequent frames* (e.g., "these × are") yield category formation by their frequent co-occurrence with intervening content words and constitute the basis for the creation of grammatical categories (Mintz, 2003). Gómez and Maye (2005) showed that 15- and 18-month-old succeed when frames have high variability in the intervening word inside the frame but failed with low variability which is in agreement with the frame-based categorization proposed by Mintz (2003). Similarly, increasing the variability of the irrelevant intervening information makes adjacent relations less statistically informative and the non-adjacent dependency more prominent allowing learners to focus on the relevant and reliable relationship among non-adjacent elements (Gómez, 2002). Interestingly these different studies converge in a similar age between 15 and 18 months old as Gerken et al. (2011) where these authors found that infants use selective attention to focus on languages having learnable grammatical patterns. These studies converge to the parallelism between the development of non-adjacent dependency learning, category formation and endogenous control of attention in the second year of life. The importance of correct tuning in attention for the acquisition of non-adjacent rules is also seen in the Lany and Gómez (2008) study, where infants younger than those of the previous studies were able to track non-adjacent dependencies if the correct attention focus was guided by training them first on the dependencies between categories. Infants later discriminated grammatical and ungrammatical items involving non-adjacent dependencies with the same category words.

Thus, the overall pattern in agreement with the progressive ability to orient attention endogenously and the close collaboration between exogenous and endogenous attention. Early on, infants need more concurrent exogenous cues such as high degree of similarity, same identity between the dependent pairs (Creel et al., 2004; Onnis et al., 2005) or prior exposure to them (Lany and Gómez, 2008; Lai and Poletiek, 2011) (for a review, see Perruchet et al., 2012), to help them to orient their attention to the relevant information (Pacton and Perruchet, 2008; Pacton et al., 2015), allowing a greater interaction between exogenous and endogenous attention. At later stages of development, after the first year of age, the improved endogenous system allows infants to rely less on the availability of these salient features to orient their attention to the relevant information.

An important point to consider is that signs of discrimination of more complex non-adjacent dependencies at a very early age have been observed only in electrophysiological studies

(Mueller et al., 2012). Online EEG measures may not reflect the same knowledge as more overt behavioral responses that require greater explicit knowledge. In that sense, it is worth considering that indicators of prediction present from birth are reflected in mismatched responses in the EEG at the presentation of unexpected events, and these early online effects reflect these more automatic prediction mechanisms. However, recent research has shown that electrophysiological indexes of conscious access, equivalent to the P300 in adults, that show a non-linear pattern, can only be tracked clearly at the end of the first year. This response associated to consciousness was visible and sustained from 12 to 15 months of age (750 ms) and may serve to amplify the sensory input through selective attention (Kouider et al., 2013). Conscious access before the first year of age may not be possible because even if the structural architecture is in place, its immaturity may not allow an adequate flux of information for conscious availability (Dehaene-Lambertz and Spelke, 2015). From the perspective presented here, this conscious access may be required for these predictions to reach a long-lasting representation that may allow the infant to show behavioral effects. More studies are needed to examine early computation of different types of non-adjacent dependencies in infancy. New research should take into account the role that variables attracting attention may have in their acquisition in order to understand when the capacity actually arises in development.

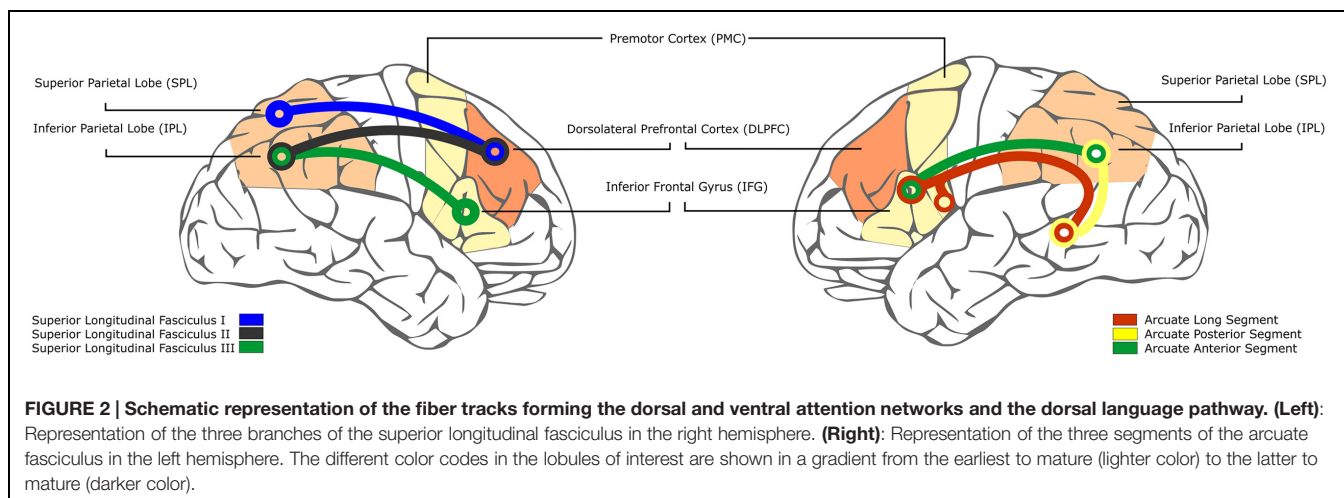# BRAIN DEVELOPMENT OF THE ATTENTION AND LANGUAGE NETWORKS

In terms of brain development, the parallel maturation of the attention and language network is also evident. This is in part unavoidable given the partial overlap between those two networks. As we have previously mentioned, a fronto-parietal network with either more ventral or dorsal distribution is related to stimulus-driven and goal-directed attention mechanisms, respectively. These areas are connected through the superior longitudinal fascicle (SLF), and the ventral and dorsal connectivity is ensured through the SLF III branch and I branch of this fascicle, respectively; these two connections have been proposed to interact through the SLF II, which connects the dorsal regions of the PFC to the ventral regions of the parietal lobe (de Schotten et al., 2011; **Figure 2**, left).

For language, a division in ventral and dorsal pathways ensures audio-motor integration and language production dorsally and language comprehension and semantic processing ventrally (Hickok and Poeppel, 2007). Direct connections between the language-related areas in the left frontal and temporal cortices are sustained dorsally through the arcuate fasciculus (AF) and ventrally by paths running through the extreme capsule (Saur et al., 2008; Brauer et al., 2011, 2013). The dorsal connection is also assured indirectly through the parietal lobe with shorter segments (Catani et al., 2005): an anterior segment connecting the premotor and inferior frontal regions with the IPL and a posterior segment connecting inferior

parietal and temporal cortices (**Figure 2**, right). There is some controversy concerning the terminations of the AF (Dick and Tremblay, 2012, for a review). Interestingly, this bundle overlaps with the SLF III, previously mentioned in relation to the ventral attention network, the anterior segment of the AF and the SLF III having a greater right lateralization (Catani et al., 2005; López-Barroso et al., 2013). Although the same nomenclature is used for attention and language in terms of ventral and dorsal streams, only the ventral attention and dorsal language stream overlap (see **Figure 2**). Although we have based this section on models of attention based on visual attention, Corbetta et al. (2008) did mention that the ventral attention network responds to different modalities. The overlap between the ventral attention and dorsal language networks is even greater if we consider that the temporal attention network shows a greater left functional lateralization, pointing again to the importance of temporal attention in speech processing (Coull et al., 2011).

During development, studies of whole ventral language connections demonstrated that newborns exhibit an adult-like ventral connection between the frontal and temporal lobes, and even children at 7 years of age have a preferential use of this pathway for sentence comprehension, in contrast to adults, who preferentially use the dorsal pathway (Brauer et al., 2011). In contrast, the dorsal pathway follows different developmental courses, with two subparts maturing at different rates. Whereas the dorsal connections reaching the premotor cortex are functional at birth, the terminations of the dorsal pathway reaching the posterior portion of Broca's area (BA 44) are still underdeveloped (Perani et al., 2011) and are not fully myelinated at the age of seven (Brauer et al., 2011, 2013). Adult studies on the learning of new languages indicate that whereas the audio-motor subpart is related to word learning (López-Barroso et al., 2013), the processing of non-adjacent elements relies on the latter subpart running from frontal BA 44 (Friederici, 2011). This pathway, which may support hierarchical (non-local) dependencies (Boeckx et al., 2014), follows a later and slower rate of development, similarly to non-adjacent, dependency rule learning.

Rapid changes are observed during the first year of life in terms of maturation. Sensory and motor systems myelinate earlier than brain systems serving higher level functions (Flechsig, 1920). Myelination starts at different times and occurs at different rates in different areas. At birth, there is little differentiation between gray and white matter in cortical areas. The primary visual cortex rapidly matures during the first 3 months with parallel myelination of optical radiations, whereas the primary auditory cortex and acoustic radiations extend over the first 3 years of life. The frontal areas and cortico-cortical connections continue to mature until puberty, but myelination is already observed during the first year in all associative regions. Diffusion measures increase with the compactness/myelination of the tracts in the left lower part of the cortico-spinal tract and in the parietal part of the AF relative to the right during the first 3 months of life. During this period the maturation of the right hemisphere is generally faster than the left (i.e., superior temporal sulcus, STS), but the inferior frontal gyrus (IFG) shows earlier left than right development. The left AF matures faster than the right

FIGURE 2 | Schematic representation of the fiber tracks forming the dorsal and ventral attention networks and the dorsal language pathway. (Left): Representation of the three branches of the superior longitudinal fasciculus in the right hemisphere. **(Right)**: Representation of the three segments of the arcuate fasciculus in the left hemisphere. The different color codes in the lobules of interest are shown in a gradient from the earliest to mature (lighter color) to the latter to mature (darker color).

and correlates with the maturation of BA 44 and the posterior part of the STS (Dehaene-Lambertz and Spelke, 2015 for a review).

Postnatal maturation shows subcortical white matter expansion in the connections to the frontal, anterior temporal, and parietal cortices, as measured by diffusion imaging and volume expansion (Hill et al., 2010). Surface expansion reflects an underlying change in synaptogenesis, dendritic arborization, gliogenesis, and intracortical myelination. The lateral temporal and parietal lobes and the dorsal and medial prefrontal regions are functionally and structurally not mature at birth. They show high expansion in cortical folding in both hemispheres in infants compared with adults. The latest maturation in synaptic density, peak cortical thickness, and mature values of gray matter density are reached in the dorsolateral prefrontal cortex. The comparison between human and macaque monkey cortices reveals that these dorsal, medial frontal and lateral parietal cortices show correlated high postnatal and evolutionary expansion (Hill et al., 2010). This pattern suggests similar patterns of cortical expansion in the development and evolution of these areas, which points to the importance of these areas for human specific functions.

These changes in connectivity at the structural level are also reflected in functional connectivity. Graph-theoretic measures of infants' brains (Power et al., 2010) indicate that the developing functional networks are in some respects similar to adult networks. The necessary connections are present; however, the brain connectivity compared to adults tends to have strong resting state functional connectivity MRI (rs-fcMRI) signal correlations with nearby regions, even during childhood. The progressively local correlations tend to weaken, whereas correlations with more distant regions, such as those between the frontal and parietal cortices, tend to increase. This trend stems from synaptic pruning that contributes to reduced local rs-fcMRI correlation, and myelination that could facilitate increased long-range connectivity.

Considering the overall attention and language networks, brain regions and connections of overlap are observed between the ventral attention network and the dorsal language network

(**Figure 2**). Ventral prefrontal and insular regions integrating the ventral attention network and the anterior segment of the AF show an early availability, whereas the IPL and their connections show a later and more progressive development. This delayed development also affects the dorsal attention network with the dorsal prefrontal regions having a slow maturation extending to childhood and with delayed maturation of the parietal lobe (Casey et al., 2000; Fuster, 2001). There is evidence showing that the left IFG is engaged in the extraction of TPs (Karuza et al., 2013) as well as the PMC (Cunillera et al., 2009) when no other cue is available to segment speech (McNealy et al., 2006; Scott-Van Zeeland et al., 2010). The early functionality of the left IFG and the premotor cortex (PMC) allows early use of TPs and stimulus-driven attention to orient to salient prosodic information and to segment speech. When the dorsal prefrontal cortex starts to be maturationally functional during the second year of life (Colombo and Cheatham, 2006), the dorsal fronto-parietal network allows for more proficient control of attention. The later maturation of the dorsal prefrontal cortex (DLPFC) and part of the ventral attention network (i.e., IPL), including the temporo-parietal junction (TPJ) (see **Figure 2**), allows progressively to (i) orient the ventral attention network to task-relevant representations (e.g., phonemes of the native language and segmented words) created in the earlier stages of development, (ii) recruit goal-directed attention, (iii) optimal functioning of the attention system, that requires the effective interaction between the two networks through the TPJ and the DLPFC (Corbetta and Shulman, 2002; Corbetta et al., 2008), necessary to accurately and selectively attend to specific stimuli and shift the focus of attention when relevant stimulation appears.

ATTENTION DEFICITS AND LANGUAGE DEVELOPMENT DISORDERS

The proposal delineated here makes a straightforward prediction in relation to the effects of attention deficits in language development. If control of attention is a function used for the

optimal acquisition of non-adjacent rule dependencies, then impairments in the development of this function should interfere with the acquisition of these rules. In contrast, early language development relying on more automatic attention mechanisms should not be affected.

Commonly, children acquire language rapidly and effortlessly. However, some children show problems acquiring language. Specific language impairment (SLI) is classically defined as a developmental disorder of language characterized by difficulty in acquiring language in the absence of neurological damage, hearing deficits, or intellectual disabilities (Bishop, 1992; Leonard, 1998). The prevalence of SLI in pre-school children is approximately 7% (Tomblin et al., 1997; Law et al., 2000). Longitudinal studies reveal that more than 70% of diagnosed cases of SLI in kindergarten persist into adulthood (Johnson et al., 1999). SLI children have been shown to have difficulties in the acquisition of non-adjacent dependencies (Hsu et al., 2014) and in the use of prosodic information for syntactic processing (Sabisch et al., 2009). In a longitudinal study, impaired prosodic processing of word stress during early development was shown to be an early marker of risk for SLI (Weber et al., 2005).

Linguistic impairments often co-occur with non-linguistic deficits, including attention-deficit/hyperactivity disorder (ADHD). Both SLI and ADHD frequently overlap within the same children, that is, comorbidity between the two disorders is commonly found (Baker and Cantwell, 1992; Benasich et al., 1993; Coster et al., 1999; Noterdaeme and Amorosa, 1999; Tomblin et al., 2000; Lindsay et al., 2007). ADHD is the most frequent diagnosis among children with language impairments (Cohen et al., 2000). Longitudinal studies suggest that SLI children have a profound risk for ADHD (Baker and Cantwell, 1987; Cantwell and Baker, 1987; Beitchman et al., 1989; Benasich et al., 1993; Redmond and Rice, 1998, 2002). More precisely, deficits in selective attention (Stevens et al., 2008) and sustained attention (Spaulding et al., 2008; Finneran et al., 2009) have been found in children with SLI. ADHD is a common childhood disorder characterized by a persistent pattern of inattention and/or developmentally inappropriate levels of hyperactivity/impulsivity (American Psychiatric Association, 2000). ADHD prevalence is approximately 10% in children (Faraone et al., 2003; Pastor and Reuben, 2008). As with SLI, children with ADHD are a highly heterogeneous group. ADHD is commonly divided into three subtypes: ADHD-Inattentive (ADHD-I), ADHD-Hyperactive-Impulsive (ADHD-H/I), and ADHD-Combined type (ADHD-C). Whereas children in the ADHD-I subgroup usually show difficulties with attention control, sustained attention and are often inattentive, ADHD-H children exhibit high levels of activity and poor impulse control. ADHD-I children do poorly in tasks requiring sustained attention, covert shifting of attention and selective attention. Thus, individual differences in the control of selective attention in infancy may be related to ADHD-I outcomes. Children in the ADHD-I group are more probable to meet criteria for learning disability than ADHD-H children (Willcutt and Pennington, 2000).

Similar to the findings on attention deficits found in SLI children, a similar pattern is present in ADHD children. Between 50 and 90% of children with ADHD have co-occurring language difficulties (Gualtieri et al., 1983; Camarata et al., 1988; Love and Thompson, 1988; Tirosh and Cohen, 1998). However, the overlap between these disorders shows an asymmetrical pattern, that is, more ADHD children have co-occurring SLI than SLI children have co-occurring ADHD (Tannock and Schachar, 1996). Higher order cognitive functions (e.g., executive functions, working memory, and attention) have been explored as possible causal deficits for SLI and ADHD disorders (e.g., Cardy et al., 2010; Hutchinson et al., 2012).

In SLI, abnormal diffusion measures are observed systematically in the SLF and AF (Verhoeven et al., 2012; Roberts et al., 2014). A more recent study showed also differences in the ventral language network (i.e., the inferior fronto-occipital fasciculus, IFOF) (Vydrova et al., 2015). The discrepancies between studies may stem from the heterogeneity of the disease with children with more semantico-pragmatic profiles that are more likely to show differences in the IFOF function and those with and without associated ADHD, which may cause an associated SLF abnormality in addition to the AF. The brain structures supporting cognitive functions commonly associated with ADHD have also been investigated. Gross anatomical changes in brain dimensions are often associated with ADHD, specifically, reduced dimensions of the caudate nucleus, the prefrontal cortex, the corpus callosum, and the cerebellar vermis (see Bush et al., 2005 for a review) and in the parietal lobes (Sowell et al., 2003) are found in ADHD. Evidence from pathophysiology research has shown that ADHD physiology involves dopaminergic and noradrenergic pathway dysfunction in the prefrontal cortex and subcortical regions of the brain (Barkley et al., 1992; Castellanos et al., 1996; Faraone and Biederman, 1998; Konrad et al., 2006). This network partially overlaps with both goal-directed attention and temporal processing. DA dysfunction affects mainly the dorsal regions of the PFC, which are those required for goal-directed attention. The subcortical regions affected (i.e., striatum) and the cerebellum are important structures for temporal processing (Coull et al., 2011).

Recent studies have provided the first evidence that temporal selective attention during speech perception predicts language outcome in preschool children. Children who selectively allocate attention to informative moments during speech, such as word onsets, demonstrate better metalinguistic capacity (Astheimer et al., 2014).

CONCLUSION

Infants acquire language exceptionally fast and without any given instruction. But, how can infants so easily achieve such a remarkable landmark, whereas adults struggle to do so? Following Kuhl's view (Kuhl, 2004), understanding how the early brain is committed to the statistical and prosodic patterns experienced early in life helps to explain the long-standing puzzle of why infants are better language learners than adults. One of the possible answers is the way their cognitive

development is structured, with functions, such as attention, appearing in an incremental fashion and assisting language learning.

Based on the characteristics of the developmental trajectory of the attention and language systems, we have outlined the hypothesis that attention development, characterized by an initial phase when attention is stimulus-driven, followed by a progressive ability to endogenously control the focus of attention, shapes the developmental trajectory of language. In the evidence reviewed here, we have seen that the learning trajectory of two types of linguistic learning (words and rules) shows a different profile in infant language development. Whereas words in fluent speech are already segmented and extracted at early stages, non-adjacent dependencies occurring over temporally distant elements are learned many months later (see **Figure 1**).

More precisely, the early segmentation and word learning abilities is profoundly influenced by the salient characteristics of the speech signal, with an important role of prosodic information. Later acquisition of more complex information associated with the extraction of more distant dependencies is influenced by variables that help infants to focus attention on the relevant elements carrying the dependency and to disregard the information that is not relevant for the acquisition of the dependencies. This trajectory goes hand in hand with the development of the ability to progressively orient attention endogenously. Early in this phase, infants require more concurrent salient cues, such as phonological similarity or identity repetition, to help them to orient their attention to the relevant information. A greater development of goal-directed attention allows infants to learn less salient, non-adjacent dependencies by relying more on endogenous cues. In terms of brain development, whereas the initial stages of development rely on the availability of some areas of the ventral attention network, including the ventral prefrontal regions and the premotor cortex, the latter stages require the maturation of more dorsal prefrontal and parietal regions (see **Figure 2**).

We consider that this development of attention in different stages allows for an earlier simplification of learning. This early learning is driven by the automatic capture of attention, creating the first building blocks that learning can lean on when control of attention allows for the extraction of more complex relations between non-adjacent elements in speech. Data from adults show that they can track both adjacent and non-adjacent information at the same time, and one information can interfere with the other (Romberg and Saffran, 2013). Thus, the inability to reorient attention away from the automatic attractors of attention is valuable in the early stages of acquisition, allowing for incremental learning.

Moreover, the same exogenous system that allows young infants to extract words using salient cues may also help them to extract complex rules. Young infants are able to succeed in non-adjacent learning that otherwise would not be available after the first year of life. In these early stages, this success of non-adjacent dependency tracking occurs only under certain conditions. Applying our present proposal to this developmental scenario, two main conditions should be fulfilled to extract non-adjacent dependencies in the early stages of development: (1) a rudimentary mechanism of endogenous attention should be available to select certain predictive elements and to disregard irrelevant information, and (2) stimulus-driven factors should be present in the linguistic input (e.g., certain degree of similarity or saliency) to automatically capture the exogenous attention system.

The implications of our hypothesis are clear in terms of the parallelism between the development of the endogenous attention system and the rule learning abilities in healthy infants. This relation is seen not only in healthy development but also in the effects of attention deficits in relation with impairments in language development. The importance of being able to exploit the available information given by exogenous cues, such as prosodic information, to orient attention endogenously is crucial not only in infant healthy development but also in studies with different pathologies.

Comprehending the cognitive processes involved in language development is of critical importance for our understanding of why, under certain conditions, language development impairment occurs. However, research in the field of language development often offers limited explanations bounded within the language domain, ignoring the importance of other cognitive functions. The present proposal overcomes these limits and presents an integrative approach to understand the role of attentional tuning during language acquisition. By reviewing the main stages of attention and language development and possible impairments, we have strengthened the importance of taking an interdisciplinary approach to the study of human development. We believe that this integrative approach exploring the role of temporal attention as a scaffold for language development can lead to a wider scope than previous proposals, allowing the development of a precise model of language and cognitive function interaction during learning that has important clinical and developmental consequences, hence providing an important contribution to the language learning and language rehabilitation fields.

AUTHOR CONTRIBUTIONS

RD-B, AM-A provided the ideas and wrote the article. FP contributed to discussions, writing and checked that the infant research review was accurate.

ACKNOWLEDGMENTS

REFERENCES

Adler, S. A., Haith, M. M., Arehart, D. M., and Lanthier, E. C. (2008). Infants' visual expectations and the processing of time. *J. Cogn. Dev.* 9, 1–25. doi: 10.1080/15248370701836568

Althaus, N., and Mareschal, D. (2012). Using saliency maps to separate competing processes in infant visual cognition. *Child Dev.* 83, 1122–1128. doi: 10.1111/j.1467-8624.2012.01766.x

American Psychiatric Association (2000). *Diagnostic and Statistical Manual of Mental Disorders*, 4th Edn. Washington, DC: American Psychiatric Association.

Astheimer, L., Janus, M., Moreno, S., and Bialystok, E. (2014). Electrophysiological measures of attention during speech perception predict metalinguistic skills in children. *Dev. Cogn. Neurosci.* 7, 1–12. doi: 10.1016/j.dcn.2013.10.005

Baker, L., and Cantwell, D. P. (1987). A prospective psychiatric follow-up of children with speech/language disorders. *J. Am. Acad. Child Adolesc. Psychiatry* 26, 546–553. doi: 10.1097/00004583-198707000-00015

Baker, L., and Cantwell, D. P. (1992). Attention deficit disorder and speech/language disorders. *Compr. Ment. Health Care* 2, 3–16.

Baldwin, D. A. (1991). Infants' contribution to the achievement of joint reference. *Child Dev.* 62, 875–890. doi: 10.2307/1131140

Baldwin, D. A. (1993). Infants' ability to consult the speaker for clues to word reference. *J. Child Lang.* 20, 395–418. doi: 10.1017/S0305000900008345

Barkley, R. A., Grodzinsky, G., and DuPaul, G. J. (1992). Frontal lobe functions in attention deficit disorder with and without hyperactivity: a review and research report. *J. Abnorm. Child Psychol.* 20, 163–188. doi: 10.1007/BF00916547

Barnes, R., and Jones, M. R. (2000). Expectancy, attention, and time. *Cogn. Psychol.* 41, 254–311. doi: 10.1006/cogp.2000.0738

Beitchman, J. H., Hood, J., Rochon, J., and Peterson, M. (1989). Empirical classification of speech/language impairment in children. II. Behavioral characteristics. *J. Am. Acad. Child Adolesc. Psychiatry* 28, 118–123. doi: 10.1097/00004583-198901000-00022

Benasich, A. A., Curtiss, S., and Tallal, P. (1993). Language, learning, and behavioral disturbances in childhood: a longitudinal perspective. *J. Am. Acad. Child Adolesc. Psychiatry* 32, 585–594. doi: 10.1097/00004583-199305000-00015

Bishop, D. V. (1992). The underlying nature of specific language impairment. *J. Child Psychol. Psychiatry* 33, 3–66. doi: 10.1111/j.1469-7610.1992.tb00858.x

Boeckx, C., Martinez-Alvarez, A., and Leivada, E. (2014). The functional neuroanatomy of serial order in language. *J. Neurolinguistics* 32, 1–15. doi: 10.1016/j.jneuroling.2014.07.001

Brannon, E. M., Libertus, M. E., Meck, W. H., and Woldorff, M. G. (2008). Electrophysiological measures of time processing in infant and adult brains: Weber's Law holds. *J. Cogn. Neurosci.* 20, 193–203. doi: 10.1162/jocn.2008.20016

Brannon, E. M., Roussel, L. W., Meck, W. H., and Woldorff, M. (2004). Timing in the baby brain. *Brain Res. Cogn. Brain Res.* 21, 227–233. doi: 10.1016/j.cogbrainres.2004.04.007

Brauer, J., Anwander, A., and Friederici, A. D. (2011). Neuroanatomical prerequisites for language functions in the maturing brain. *Cereb. Cortex* 21, 459–466. doi: 10.1093/cercor/bhq108

Brauer, J., Anwander, A., Perani, D., and Friederici, A. D. (2013). Dorsal and ventral pathways in language development. *Brain Lang.* 127, 289–295. doi: 10.1016/j.bandl.2013.03.001

Brooks, R., and Meltzoff, A. N. (2002). NIH public access. *Dev. Psychol.* 38, 958–966. doi: 10.1016/j.micinf.2011.07.011.Innate

Bruner, J. (1983). *Child's Talk: Learning to Use Language.* New York: Academic Press.

Bush, G., Valera, E. M., and Seidman, L. J. (2005). Functional neuroimaging of attention-deficit/hyperactivity disorder: a review and suggested future directions. *Biol. Psychiatry* 57, 1273–1284. doi: 10.1016/j.biopsych.2005.01.034

Camarata, S. M., Hughes, C. A., and Ruhl, K. L. (1988). Mild/moderate behaviorally disordered students: a population at risk for language disorders. *Lang. Speech Hear. Serv. Sch.* 19, 191–200. doi: 10.1044/0161-1461.1902.191

Cantwell, D. P., and Baker, L. (1987). Prevalence and type of psychiatric disorder and developmental disorders in three speech and language groups. *J. Commun. Disord.* 20, 151–160. doi: 10.1016/0021-9924(87)90006-2

Cardy, J. E. O., Tannock, R., Johnson, A. M., and Johnson, C. J. (2010). The contribution of processing impairments to SLI: insights from

attention-deficit/hyperactivity disorder. *J. Commun. Disord.* 43, 77–91. doi: 10.1016/j.jcomdis.2009.09.003

Carpenter, M., Nagell, K., and Tomasello, M. (1998). Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monogr. Soc. Res. Child Dev.* 63, i–vi, 1–143. doi: 10.2307/1166214

Casey, B. J., Davidson, M. C., Hara, Y., Thomas, K. M., Martinez, A., Galvan, A., et al. (2004). Early development of subcortical regions involved in non-cued attention switching. *Dev. Sci.* 7, 534–542. doi: 10.1111/j.1467-7687.2004.00377.x

Casey, B. J., Giedd, J. N., and Thomas, K. M. (2000). Structural and functional brain development and its relation to cognitive development. *Biol. Psychol.* 54, 241–257. doi: 10.1016/S0301-0511(00)00058-2

Castellanos, F. X., Giedd, J. N., Marsh, W. L., Hamburger, S. D., Vaituzis, A. C., Dickstein, D. P., et al. (1996). Quantitative brain magnetic resonance imaging in attention-deficit hyperactivity disorder. *Arch. Gen. Psychiatry* 53, 607–616. doi: 10.1001/archpsyc.1996.01830070053009

Catani, M., Jones, D. K., and Ffytche, D. H. (2005). Perisylvian language networks of the human brain. *Ann. Neurol.* 57, 8–16. doi: 10.1002/ana.20319

Chica, A. B., Bartolomeo, P., and Lupiáñez, J. (2013). Two cognitive and neural systems for endogenous and exogenous spatial attention. *Behav. Brain Res.* 237, 107–123. doi: 10.1016/j.bbr.2012.09.027

Chomsky, N. (1957). *Syntactic Structures.* The Hague/Paris: Mouton.

Christophe, A., Millotte, S., Bernal, S., and Lidz, J. (2008). Bootstrapping lexical and syntactic acquisition. *Lang. Speech* 51, 61–75. doi: 10.1177/00238309080510010501

Cohen, N. J., Vallance, D. D., Barwick, M., Im, N., Menna, R., Horodezky, N. B., et al. (2000). The interface between ADHD and language impairment: an examination of language, achievement, and cognitive processing. *J. Child Psychol. Psychiatry* 41, 353–362. doi: 10.1111/1469-7610.00619

Colombo, J. (2001). The development of visual attention in infancy. *Annu. Rev. Psychol.* 52, 337–367. doi: 10.1146/annurev.psych.52.1.337

Colombo, J., and Cheatham, C. L. (2006). The emergence and basis of endogenous attention in infancy and early childhood. *Adv. Child Dev. Behav.* 34, 283–322. doi: 10.1016/S0065-2407(06)80010-8

Cooper, R. P., and Aslin, R. N. (1990). Preference for infant-directed speech in the first month after birth. *Child Dev.* 61, 1584–1595. doi: 10.2307/1130766

Corbetta, M., Patel, G., and Shulman, G. L. (2008). The reorienting system of the human brain: from environment to theory of mind. *Neuron* 58, 306–324. doi: 10.1016/j.neuron.2008.04.017

Corbetta, M., and Shulman, G. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nat. Rev. Neurosci.* 3, 201–215. doi: 10.1038/nrn755

Correa, Á, Lupiáñez, J., Madrid, E., and Tudela, P. (2006). Temporal attention enhances early visual processing: a review and new evidence from event-related potentials. *Brain Res.* 1076, 116–128. doi: 10.1016/j.brainres.2005.11.074

Coster, F. W., Goorhuis-Brouwer, S. M., Nakken, H., and Spelberg, H. C. (1999). Specific language impairments and behavioural problems. *Folia Phoniatr. Logop.* 51, 99–107. doi: 10.1159/000021484

Coull, J. T., Cheng, R.-K., and Meck, W. H. (2011). Neuroanatomical and neurochemical substrates of timing. *Neuropsychopharmacology* 36, 3–25. doi: 10.1038/npp.2010.113

Courage, M. L., Reynolds, G. D., and Richards, J. E. (2006). Infants' attention to patterned stimuli: developmental change from 3 to 12 months of age. *Child Dev.* 77, 680–695. doi: 10.1111/j.1467-8624.2006.00897.x

Courage, M. L., and Richards, J. E. (2008). "Attention," in *Encyclopedia of Infant and Early Childhood Development*, eds M. M. Haith and J. B. Benson (Oxford: Elsevier), 106–117.

Creel, S. C., Newport, E. L., and Aslin, R. N. (2004). Distant melodies: statistical learning of nonadjacent dependencies in tone sequences. *J. Exp. Psychol. Learn. Mem. Cogn.* 30, 1119–1130. doi: 10.1037/0278-7393.30.5.1119

Cunillera, T., Càmara, E., Toro, J. M., Marco-Pallares, J., Sebastián-Galles, N., Ortiz, H., et al. (2009). Time course and functional neuroanatomy of speech segmentation in adults. *Neuroimage* 48, 541–553. doi: 10.1016/j.neuroimage.2009.06.069

de Diego-Balaguer, R., Rodríguez-Fornells, A., and Bachoud-Lévi, A. C. (2015). Prosodic cues enhance rule learning by changing speech

segmentation mechanisms. *Front. Psychol.* 6:1478. doi: 10.3389/fpsyg.2015.01478

de Schotten, M. T., Dell'Acqua, F., Forkel, S. J., Simmons, A., Vergani, F., Murphy, D. G. M., et al. (2011). A lateralized brain network for visuospatial attention. *Nat. Neurosci.* 14, 1245–1246. doi: 10.1038/nn.2905

Dehaene, S., Meyniel, F., Wacongne, C., Wang, L., and Pallier, C. (2015). The neural representation of sequences: from transition probabilities to algebraic patterns and linguistic trees. *Neuron* 88, 2–19. doi: 10.1016/j.neuron.2015.09.019

Dehaene-Lambertz, G., and Spelke, E. S. (2015). The infancy of the human brain. *Neuron* 88, 93–109. doi: 10.1016/j.neuron.2015.09.026

Demany, L., Mckenzie, B., and Vurpillot, E. (1977). Rhythm perception in early infancy. *Nature* 266, 718–719. doi: 10.1038/266718a0

Diamond, A. (2007). Consequences of variations in genes that affect dopamine in prefrontal cortex. *Cereb. Cortex* 17, 1–19. doi: 10.1093/cercor/bhm082

Dick, A. S., and Tremblay, P. (2012). Beyond the arcuate fasciculus: consensus and controversy in the connectional anatomy of language. *Brain* 135, 3529–3550. doi: 10.1093/brain/aws222

Doherty, J. R., Rao, A., Mesulam, M. M., and Nobre, A. C. (2005). Synergistic effect of combined temporal and spatial expectations on visual attention. *J. Neurosci.* 25, 8259–8266. doi: 10.1523/JNEUROSCI.1821-05.2005

Elison, J. T., Wolff, J. J., Heimer, D. C., Paterson, S. J., Gu, H., Hazlett, H. C., et al. (2013). Frontolimbic neural circuitry at 6 months predicts individual differences in joint attention at 9 months. *Dev. Sci.* 16, 186–197. doi: 10.1111/desc.12015

Endress, A. D., Nespor, M., and Mehler, J. (2009). Perceptual and memory constraints on language acquisition. *Trends Cogn. Sci.* 13, 348–353. doi: 10.1016/j.tics.2009.05.005

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

Fair, D. A., Cohen, A. L., Power, J. D., Dosenbach, N. U. F., Church, J. A., Miezin, F. M., et al. (2009). Functional brain networks develop from a "local to distributed" organization. *PLoS Comput. Biol.* 5:e1000381. doi: 10.1371/journal.pcbi.1000381

Fair, D. A., Dosenbach, N. U. F., Church, J. A., Cohen, A. L., Brahmbhatt, S., Miezin, F. M., et al. (2007). Development of distinct control networks through segregation and integration. *Proc. Natl. Acad. Sci. U.S.A.* 104, 13507–13512. doi: 10.1073/pnas.0705843104

Faraone, S. V., and Biederman, J. (1998). Neurobiology of attention-deficit hyperactivity disorder. *Biol. Psychiatry* 44, 951–958. doi: 10.1016/S0006-3223(98)00240-6

Faraone, S. V., Sergeant, J., Gillberg, C., and Biederman, J. (2003). The worldwide prevalence of ADHD: is it an American condition? *World Psychiatry* 2, 104–113.

Fernald, A. (1985). Four-month-old infants prefer to listen to motherese. *Infant Behav. Dev.* 8, 181–195. doi: 10.1016/S0163-6383(85)80005-9

Finneran, D. A., Francis, A. L., and Leonard, L. B. (2009). Sustained attention in children with specific language impairment (SLI). *J. Speech Lang. Hear. Res.* 52, 915–929. doi: 10.1044/1092-4388(2009/07-0053)

Flechsig, P. (1920). *Anatomie des Menschlichen Gehirns und Rückenmarks auf Myelogenetishcher Grundlage.* Liepzig: George Thomas.

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., Mueller, J. L., and Oberecker, R. (2011). Precursors to natural grammar learning: preliminary evidence from 4-month-old infants. *PLoS ONE* 6:e17920. doi: 10.1371/journal.pone.0017920

Fuster, J. M. (2001). The prefrontal cortex–an update: time is of the essence. *Neuron* 30, 319–333. doi: 10.1016/S0896-6273(01)00285-9

Gao, W., Zhu, H., Giovanello, K., Smith, K., Shen, D., Gilmore, J., et al. (2009). "Emergence of the brain's default network: evidence from two-week-old to four-year-old healthy pediatric subjects," in *Proceedings of the 17th Scientific Meeting of the International Society for Magnetic Resonance in Medicine*, Honolulu, HI, 3434.

Gerken, L. (2006). Decisions, decisions: infant language learning when multiple generalizations are possible. *Cognition* 98, 67–74. doi: 10.1016/j.cognition.2005.03.003

Gerken, L., Balcomb, F. K., and Minton, J. L. (2011). Infants avoid "labouring in vain" by attending more to learnable than unlearnable linguistic patterns. *Dev. Sci.* 14, 972–979. doi: 10.1111/j.1467-7687.2011.01046.x

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

Gogtay, N., Giedd, J. N., Lusk, L., Hayashi, K. M., Greenstein, D., Vaituzis, A. C., et al. (2004). Dynamic mapping of human cortical development during childhood through early adulthood. *Proc. Natl. Acad. Sci. U.S.A.* 101, 8174–8179. doi: 10.1073/pnas.0402680101

Gómez, R., and Maye, J. (2005). The developmental trajectory of nonadjacent dependency learning. *Infancy* 7, 183–206. doi: 10.1207/s15327078in0702_4

Gómez, R. L. (2002). Variability and detection of invariant structure. *Psychol. Sci.* 13, 431–436. doi: 10.1111/1467-9280.00476

Gonzalez-Gomez, N., and Nazzi, T. (2012). Acquisition of nonadjacent phonological dependencies in the native language during the first year of life. *Infancy* 17, 498–524. doi: 10.1111/j.1532-7078.2011.00104.x

Gualtieri, C. T., Koriath, U., Van Bourgondien, M., and Saleeby, N. (1983). Language disorders in children referred for psychiatric services. *J. Am. Acad. Child Psychiatry* 22, 165–171. doi: 10.1016/S0002-7138(09)62330-X

Haith, M. M., Hazan, C., and Goodman, G. S. (1988). Expectation and anticipation of dynamic visual events by 3.5-month-old babies. *Child Dev.* 59, 467–479. doi: 10.2307/1130325

Hickok, G., and Poeppel, D. (2007). The cortical organization of speech processing. *Nat. Rev. Neurosci.* 8, 393–402. doi: 10.1038/nrn2113

Hill, J., Inder, T., Neil, J., Dierker, D., Harwell, J., and Van Essen, D. (2010). Similar patterns of cortical expansion during human development and evolution. *Proc. Natl. Acad. Sci. U.S.A.* 107, 13135–13140. doi: 10.1073/pnas.1001229107

Hood, B. M. (1995). "Shifts of visual attention in the human infant: a neuroscientific approach," in *Advances in Infancy Research*, Vol. 9, eds C. Rovee-Collier and L. P. Lipsitt (Norwood, NJ: Ablex), 163–216.

Hsu, H. J., Tomblin, J. B., and Christiansen, M. H. (2014). Impaired statistical learning of non-adjacent dependencies in adolescents with specific language impairment. *Front. Psychol.* 5:175. doi: 10.3389/fpsyg.2014.00175

Hunnius, S., and Geuze, R. H. (2004). Gaze shifting in infancy: a longitudinal study using dynamic faces and abstract stimuli. *Infant Behav. Dev.* 27, 397–416. doi: 10.1016/j.infbeh.2004.02.003

Hutchinson, E., Bavin, E., Efron, D., and Sciberras, E. (2012). A comparison of working memory profiles in school-aged children with Specific Language Impairment, Attention Deficit/Hyperactivity Disorder, Comorbid SLI and ADHD and their typically developing peers. *Child Neuropsychol.* 18, 190–207. doi: 10.1080/09297049.2011.601288

Ibbotson, P., Lieven, E. V. M., and Tomasello, M. (2013). The attention-grammar interface: eye-gaze cues structural choice in children and adults. *Cogn. Linguist.* 24, 457–481. doi: 10.1515/cog-2013-0020

Jackendoff, R. (2002). *Foundations of Language.* Oxford: Oxford University Press.

Johnson, C. J., Beitchman, J. H., Young, A., Escobar, M., Atkinson, L., Wilson, B., et al. (1999). Fourteen-year follow-up of children with and without speech/language impairments: speech/language stability and outcomes. *J. Speech Lang. Hear. Res.* 42, 744–760. doi: 10.1044/jslhr.4203.744

Johnson, E. K. (2008). Infants use prosodically conditioned acoustic-phonetic cues to extract words from speech. *J. Acoust. Soc. Am.* 123, EL144–EL148. doi: 10.1121/1.2908407

Johnson, K. A., Burrowes, E., and Coull, J. T. (2015). Children can implicitly, but not voluntarily, direct attention in time. *PLoS ONE* 10:e0123625. doi: 10.1371/journal.pone.0123625

Johnson, M. H. (1990). Cortical maturation and the development of visual attention in early infancy. *J. Cogn. Neurosci.* 2, 81–95. doi: 10.1162/jocn.1990.2.2.81

Jusczyk, P. W. (2002). How infants adapt speech-processing capacities to native-language structure. *Curr. Dir. Psychol. Sci.* 11, 15–18. doi: 10.1111/1467-8721.00159

Kanpem, A., Van Parmaksiz, G., Van De Vijver, R., and Hohle, B. (2008). "Metrical and statistical cues for word segmentation: vowel harmony and word stress as cues to word boundaries by 6-and-9-month-old turkish learnes," in *Language Acquisition and Development*, eds A. Gavarró and M. J. Freitas (Newcastle: Cambridge Scholars Publishing), 313–324.

Karuza, E. A., Newport, E. L., Aslin, R. N., Starling, S. J., Tivarus, M. E., and Bavelier, D. (2013). The neural correlates of statistical learning in a word segmentation task: an fMRI study. *Brain Lang.* 127, 46–54. doi: 10.1016/j.bandl.2012.11.007

Konrad, K., Neufang, S., Hanisch, C., Fink, G. R., and Herpertz-Dahlmann, B. (2006). Dysfunctional attentional networks in children with attention deficit/hyperactivity disorder: evidence from an event-related functional magnetic resonance imaging study. *Biol. Psychiatry* 59, 643–651. doi: 10.1016/j.biopsych.2005.08.013

Kotz, S. A., and Schwartze, M. (2010). Cortical speech processing unplugged: a timely subcortico-cortical framework. *Trends Cogn. Sci.* 14, 392–399. doi: 10.1016/j.tics.2010.06.005

Kouider, S., Stahlhut, C., Gelskov, S. V., Barbosa, L. S., Dutat, M., de Gardelle, V., et al. (2013). A neural marker of perceptual consciousness in infants. *Science* 340, 376–380. doi: 10.1126/science.1232509

Kristensen, L. B., Wang, L., Petersson, K. M., and Hagoort, P. (2012). The interface between language and attention: prosodic focus marking recruits a general attention network in spoken language comprehension. *Cereb. Cortex* 23, 1836–1848. doi: 10.1093/cercor/bhs164

Kuhl, P. K. (2004). Early language acquisition: cracking the speech code. *Nat. Rev. Neurosci.* 5, 831–843. doi: 10.1038/nrn1533

Lai, J., and Poletiek, F. H. (2011). The impact of adjacent-dependencies and staged-input on the learnability of center-embedded hierarchical structures. *Cognition* 118, 265–273. doi: 10.1016/j.cognition.2010.11.011

Landry, R., and Bryson, S. E. (2004). Impaired disengagement of attention in young children with autism. *J. Child Psychol. Psychiatry* 45, 1115–1122. doi: 10.1111/j.1469-7610.2004.00304.x

Lany, J., and Gómez, R. L. (2008). Twelve-month-old infants benefit from prior experience in statistical learning. *Psychol. Sci.* 19, 1247–1252. doi: 10.1111/j.1467-9280.2008.02233.x

Large, E. W., and Jones, M. R. (1999). The dynamics of attending: how people track time-varying events. *Psychol. Rev.* 106, 119–159. doi: 10.1037/0033-295X.106.1.119

Law, J., Boyle, J., Harris, F., Harkness, A., and Nye, C. (2000). Prevalence and natural history of primary speech and language delay: findings from a systematic review of the literature. *Int. J. Lang. Commun. Disord.* 35, 165–188. doi: 10.1080/136828200247133

Leonard, L. B. (1998). *Children with Specific Language Impairment.* Cambridge, MA: MIT Press.

Levelt, W. J. M. (1989). *Speaking: From Intention to Articulation.* Cambridge, MA: MIT Press.

Lindsay, G., Dockrell, J. E., and Strand, S. (2007). Longitudinal patterns of behaviour problems in children with specific speech and language difficulties: child and contextual factors. *Br. J. Educ. Psychol.* 77, 811–828. doi: 10.1348/000709906X171127

López-Barroso, D., Catani, M., Ripollés, P., Dell'Acqua, F., Rodríguez-Fornells, A., and de Diego-Balaguer, R. (2013). Word learning is mediated by the left arcuate fasciculus. *Proc. Natl. Acad. Sci. U.S.A.* 110, 13168–13173. doi: 10.1073/pnas.1301696110

Love, A. J., and Thompson, M. G. (1988). Language disorders and attention deficit disorders in young children referred for psychiatric services: analysis of prevalence and a conceptual synthesis. *Am. J. Orthopsychiatry* 58, 52–64. doi: 10.1111/j.1939-0025.1988.tb01566.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

Mattys, S. L., Jusczyk, P. W., Luce, P. A., and Morgan, J. L. (1999). Phonotactic and prosodic effects on word segmentation in infants. *Cogn. Psychol.* 38, 465–494. doi: 10.1006/cogp.1999.0721

McNealy, K., Mazziotta, J. C., and Dapretto, M. (2006). Cracking the language code: neural mechanisms underlying speech parsing. *J. Neurosci.* 26, 7629–7639. doi: 10.1523/JNEUROSCI.5501-05.2006

Mintz, T. H. (2003). Frequent frames as a cue for grammatical categories in child directed speech. *Cognition* 90, 91–117. doi: 10.1016/S0010-0277(03)00140-9

Mintz, T., and Walker, R. (2006). Infant's sensitivity to vowel harmony and its role in word segmentation. *Paper Presented at the Annual Meeting of the LSA,* Albuquerque, NM.

Mueller, J. L., Bahlmann, J., and Friederici, A. D. (2010). Learnability of embedded syntactic structures depends on prosodic cues. *Cogn. Sci.* 34, 338–349. doi: 10.1111/j.1551-6709.2009.01093.x

Mueller, J. L., Friederici, A. D., and Mannel, C. (2012). Auditory perception at the root of language learning. *Proc. Natl. Acad. Sci. U.S.A.* 109, 15953–15958. doi: 10.1073/pnas.1204319109

Myachykov, A., Tomblin, R. S., and Posner, M. I. (2005). Attention and empirical studies of grammar. *Linguist. Rev.* 22, 347–364. doi: 10.1515/tlir.2005.22.2-4.347

Nazzi, T., Bertoncini, J., and Bijeljac-Babic, R. (2009). A perceptual equivalent of the labial-coronal effect in the first year of life. *J. Acoust. Soc. Am.* 126, 1440–1446. doi: 10.1121/1.3158931

Nazzi, T., Bertoncini, J., and Mehler, J. (1998). Language discrimination by newborns: toward an understanding of the role of rhythm. *J. Exp. Psychol. Hum. Percept. Perform.* 24, 756–766. doi: 10.1037/0096-1523.24.3.756

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

Nobre, A., Correa, A., and Coull, J. (2007). The hazards of time. *Curr. Opin. Neurobiol.* 17, 465–470. doi: 10.1016/j.conb.2007.07.006

Noterdaeme, M., and Amorosa, H. (1999). Evaluation of emotional and behavioral problems in language impaired children using the Child Behavior Checklist. *Eur. Child Adolesc. Psychiatry* 8, 71–77. doi: 10.1007/s007870050087

Onnis, L., Monaghan, P., Richmond, K., and Chater, N. (2005). Phonology impacts segmentation in online speech processing. *J. Mem. Lang.* 53, 225–237. doi: 10.1016/j.jml.2005.02.011

Pacton, S., and Perruchet, P. (2008). An attention-based associative account of adjacent and nonadjacent dependency learning. *J. Exp. Psychol. Learn. Mem. Cogn.* 34, 80–96. doi: 10.1037/0278-7393.34.1.80

Pacton, S., Sobaco, A., and Perruchet, P. (2015). Is an attention-based associative account of adjacent and nonadjacent dependency learning valid? *Acta Psychol. (Amst)* 157, 195–199. doi: 10.1016/j.actpsy.2015.03.002

Parise, E., Friederici, A. D., and Striano, T. (2010). "Did you call me?" 5-month-old infants own name guides their attention. *PLoS ONE* 5:e14208. doi: 10.1371/journal.pone.0014208

Pastor, P. N., and Reuben, C. A. (2008). Diagnosed attention deficit hyperactivity disorder and learning disability: United States, 2004-2006. *Vital Health Stat.* 10, 1–14.

Pelucchi, B., Hay, J. F., and Saffran, J. R. (2009). Statistical learning in a natural language by 8-month-old infants. *Child Dev.* 80, 674–685. doi: 10.1111/j.1467-8624.2009.01290.x

Peña, M., Bonatti, L. L., Nespor, M., and Mehler, J. (2002). Signal-driven computations in speech processing. *Science* 298, 604–607. doi: 10.1126/science.1072901

Perani, D., Saccuman, M. C., Scifo, P., Anwander, A., Awander, A., Spada, D., et al. (2011). Neural language networks at birth. *Proc. Natl. Acad. Sci. U.S.A.* 108, 16056–16061. doi: 10.1073/pnas.1102991108

Perruchet, P., Poulin-Charronnat, B., and Pacton, S. (2012). "Learning nonadjacent dependencies," in *Encyclopedia of the Sciences of Learning*, ed. N. M. Seel (New York, NY: Springer-Verlag), 1941–1944.

Posner, M. I., and Cohen, Y. (1984). Components of visual orienting. *Atten. Perform.* 32, 531–556.

Posner, M. I., and Petersen, S. E. (1990). The attention system of the human brain. *Annu. Rev. Neurosci.* 13, 25–42. doi: 10.1146/annurev.ne.13.030190.000325

Power, J. D., Fair, D. A., Schlaggar, B. L., and Petersen, S. E. (2010). The development of human functional brain networks. *Neuron* 67, 735–748. doi: 10.1016/j.neuron.2010.08.017.The

Redmond, S. M., and Rice, M. L. (1998). The socioemotional behaviors of children with SLI: Social Adaptation or Social Deviance? *J. Speech Lang. Hear. Res.* 41, 688–700. doi: 10.1044/jslhr.4103.688

Redmond, S. M., and Rice, M. L. (2002). Stability of behavioral ratings of children with SLI. *J. Speech Lang. Hear. Res.* 45, 190–201. doi: 10.1044/1092-4388(2002/014)

Richards, J. E., Reynolds, G. D., and Courage, M. L. (2010). The neural bases of infant attention. *Curr. Dir. Psychol. Sci.* 19, 41–46. doi: 10.1177/0963721409360003

Roberts, T. P. L., Heiken, K., Zarnow, D., Dell, J., Nagae, L., Blaskey, L., et al. (2014). Left hemisphere diffusivity of the arcuate fasciculus: influences of autism spectrum disorder and language impairment. *AJNR Am. J. Neuroradiol.* 35, 587–592. doi: 10.3174/ajnr.A3754

Rohenkohl, G., Gould, I. C., Pessoa, J., and Nobre, A. C. (2014). Combining spatial and temporal expectations to improve visual perception. *J. Vis.* 14, 1–13. doi: 10.1167/14.4.8

Romberg, A. R., and Saffran, J. R. (2013). All together now: concurrent learning of multiple structures in an artificial language. *Cogn. Sci.* 37, 1290–1320. doi: 10.1111/cogs.12050

Rueda, M. R., Pozuelos, J. P., and Cómbita, L. M. (2015). Cognitive Neuroscience of Attention From brain mechanisms to individual differences in efficiency. *AIMS Neurosci.* 2, 183–202. doi: 10.3934/Neuroscience.2015.4.183

Ruff, H. A., and Rothbart, M. K. (1996). *Attention in Early Development: Themes and Variations.* New York: Oxford University Press.

Sabisch, B., Hahne, C. A. A., Glass, E., von Suchodoletz, W., and Friederici, A. D. (2009). Children with specific language impairment: the role of prosodic processes in explaining difficulties in processing syntactic information. *Brain Res.* 1261, 37–44. doi: 10.1016/j.brainres.2009.01.012

Saffran, J. R., Aslin, R. N., and Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science* 274, 1926–1928. doi: 10.1126/science.274.5294.1926

Saffran, J. R., Newport, E. L., Aslin, R. N., Tunick, R. A., and Barrueco, S. (1997). Incidental language learning: listening (and Learning) out of the Corner of Your Ear. *Psychol. Sci.* 8, 101–106. doi: 10.1111/j.1467-9280.1997.tb00690.x

Sanabria, D., Capizzi, M., and Correa, A. (2011). Rhythms that speed you up. *J. Exp. Psychol. Hum. Percept. Perform.* 37, 236–244. doi: 10.1037/a0019956

Sanders, L. D., Newport, E. L., and Neville, H. J. (2002). Segmenting nonsense: an event-related potential index of perceived onsets in continuous speech. *Nat. Neurosci.* 5, 700–703. doi: 10.1038/nn873

Sandoval, M., and Gómez, R. L. (2013). The development of nonadjacent dependency learning in natural and artificial languages. *Wiley Interdisc. Rev. Cogn. Sci.* 4, 511–522. doi: 10.1002/wcs.1244

Santelmann, L. M., and Jusczyk, P. W. (1998). Sensitivity to discontinuous dependencies in language learners: evidence for limitations in processing space. *Cognition* 69, 105–134. doi: 10.1016/S0010-0277(98)00060-2

Saur, D., Kreher, B. W., Schnell, S., Kummerer, D., Kellmeyer, P., Vry, M. S., et al. (2008). Ventral and dorsal pathways for language. *Proc. Natl. Acad. Sci. U.S.A.* 105, 18035–18040. doi: 10.1073/pnas.0805234105

Scott-Van Zeeland, A. A., Mcnealy, K., Wang, A. T., Sigman, M., Bookheimer, S. Y., and Dapretto, M. (2010). no neural evidence of statistical learning during exposure to artificial languages in children with autism spectrum disorders ashley. *Biol. Psychiatry* 68, 345–351. doi: 10.1016/j.biopsych.2010.01.011.No

Seidl, A., and Johnson, E. K. (2008). Boundary alignment enables 11-month-olds to segment vowel initial words from speech. *J. Child Lang.* 35, 1–24. doi: 10.1017/S0305000907008215

Smith, L. B., Yu, C., and Pereira, A. F. (2011). Not your mother's view: the dynamics of toddler visual experience. *Dev. Sci.* 14, 9–17. doi: 10.1111/j.1467-7687.2009.00947.x.Not

Soderstrom, M., Wexler, K., and Jusczyk, P. (2002). "English-learning toddlers' sensitivity to agreement morphology in receptive grammar," in *Proceedings of the 26th Annual Boston University Conference on Language Development,* eds B. Skarabela, S. A. Fish, and A. H. J. Do (Somerville, MA: Cascadilla Press), 643–652.

Sowell, E. R., Thompson, P. M., Welcome, S. E., Henkenius, A. L., Toga, A. W., and Peterson, B. S. (2003). Cortical abnormalities in children and adolescents with attention-deficit hyperactivity disorder. *Lancet* 362, 1699–1707. doi: 10.1016/S0140-6736(03)14842-8

Spaulding, T. J., Plante, E., and Vance, R. (2008). Sustained selective attention skills of preschool children with specific language impairment: evidence for separate attentional capacities. *J. Speech Lang. Hear. Res.* 51, 16–34. doi: 10.1044/1092-4388(2008/002)

Stechler, G., and Latz, E. (1966). Some observations on attention and arousal in the human infant. *J. Am. Acad. Child Psychiatry* 5, 517–525. doi: 10.1016/S0002-7138(09)62098-7

Stevens, C., Fanning, J., Coch, D., Sanders, L., and Neville, H. (2008). Neural mechanisms of selective auditory attention are enhanced by computerized training: electrophysiological evidence from language-impaired and typically developing children. *Brain Res.* 1205, 55–69. doi: 10.1016/j.brainres.2007.10.108

Tannock, R., and Schachar, R. (1996). "Executive dysfunction as an underlying mechanism of behavior and language problems in attention deficit hyperactivity disorder," in *Language, Learning, and Behavior Disorders: Developmental, Biological, and Clinical Perspectives,* eds J. H. Beitchman, N. Cohen, M. Konstantareas, and R. Tannock (New York: Cambridge University Press), 128–155.

Teinonen, T., Fellman, V., Näätänen, R., Alku, P., and Huotilainen, M. (2009). Statistical language learning in neonates revealed by event-related brain potentials. *BMC Neurosci.* 10:21. doi: 10.1186/1471-2202-10-21

Thiessen, E. D., Hill, E. A., and Saffran, J. R. (2005). Infant-directed speech facilitates word segmentation. *Infancy* 7, 53–71. doi: 10.1207/s15327078in0701_5

Tirosh, E., and Cohen, A. (1998). Language deficit with attention-deficit disorder: a prevalent comorbidity. *J. Child Neurol.* 13, 493–497. doi: 10.1177/088307389801301005

Tomblin, J. B., Records, N. L., Buckwalter, P., Zhang, X., Smith, E., and O'Brien, M. (1997). Prevalence of specific language impairment in kindergarten children. *J. Speech Lang. Hear. Res.* 40, 1245–1260. doi: 10.1044/jslhr.4006.1245

Tomblin, J. B., Zhang, X., Buckwalter, P., and Catts, H. (2000). The association of reading disability, behavioral disorders, and language impairment among second-grade children. *J. Child Psychol. Psychiatry* 41, 473–482. doi: 10.1111/1469-7610.00632

Tomblin, R. S., and Myachykov, A. (2015). "Attention and salience," in *Mouton Handbook of Cognitive Linguistics,* eds E. Dąbrowska and D. Divjak (Berlin: de Gruyter Mouton).

Toro, J. M., Sinnett, S., and Soto-Faraco, S. (2005). Speech segmentation by statistical learning depends on attention. *Cognition* 97, B25–B34. doi: 10.1016/j.cognition.2005.01.006

Toro, J. M., and Trobalón, J. B. (2005). Statistical computations over a speech stream in a rodent. *Percept. Psychophys.* 67, 867–875. doi: 10.3758/BF03193539

Trainor, L. J., and Desjardins, R. N. (2002). Pitch characteristics of infant-directed speech affect infants' ability to discriminate vowels. *Psychon. Bull. Rev.* 9, 335–340. doi: 10.3758/BF03196290

Tummeltshammer, K. S., Mareschal, D., and Kirkham, N. Z. (2014). Infants' selective attention to reliable visual cues in the presence of salient distractors. *Child Dev.* 85, 1981–1994. doi: 10.1111/cdev.12239

Turk-Browne, N. B., Jungé, J., and Scholl, B. J. (2005). The automaticity of visual statistical learning. *J. Exp. Psychol. Gen.* 134, 552–564. doi: 10.1037/0096-3445.134.4.552

Uylings, H. B. M. (2006). Development of the human cortex and the concept of "Critical" or "Sensitive" Periods. *Lang. Learn.* 56, 59–90. doi: 10.1111/j.1467-9922.2006.00355.x

vanMarle, K., and Wynn, K. (2006). Six-month-old infants use analog magnitudes to represent duration. *Dev. Sci.* 9, F41–F49. doi: 10.1111/j.1467-7687.2006.00508.x

Verhoeven, J. S., Rommel, N., Prodi, E., Leemans, A., Zink, I., Vandewalle, E., et al. (2012). Is there a common neuroanatomical substrate of language deficit between autism spectrum disorder and specific language impairment? *Cereb. Cortex* 22, 2263–2271. doi: 10.1093/cercor/bhr292

Vydrova, R., Komarek, V., Sanda, J., Sterbova, K., Jahodova, A., Maulisova, A., et al. (2015). Structural alterations of the language connectome in children with specific language impairment. *Brain Lang.* 151, 35–41. doi: 10.1016/j.bandl.2015.10.003

Weber, C., Hahne, A., Friedrich, M., and Friederici, A. D. (2005). Reduced stress pattern discrimination in 5-month-olds as a marker of risk for later language impairment: neurophysiologial evidence. *Brain Res. Cogn. Brain Res.* 25, 180–187. doi: 10.1016/j.cogbrainres.2005.05.007

Willcutt, E. G., and Pennington, B. F. (2000). Comorbidity of reading disability and attention-deficit/hyperactivity disorder: differences by gender and subtype. *J. Learn. Disabil.* 33, 179–191. doi: 10.1177/002221940003300206

Wolff, P. H. (1965). THE DEVELOPMENT OF ATTENTION IN YOUNG INFANTS. *Ann. N. Y. Acad. Sci.* 118, 815–830. doi: 10.1111/j.1749-6632.1965.tb40153.x