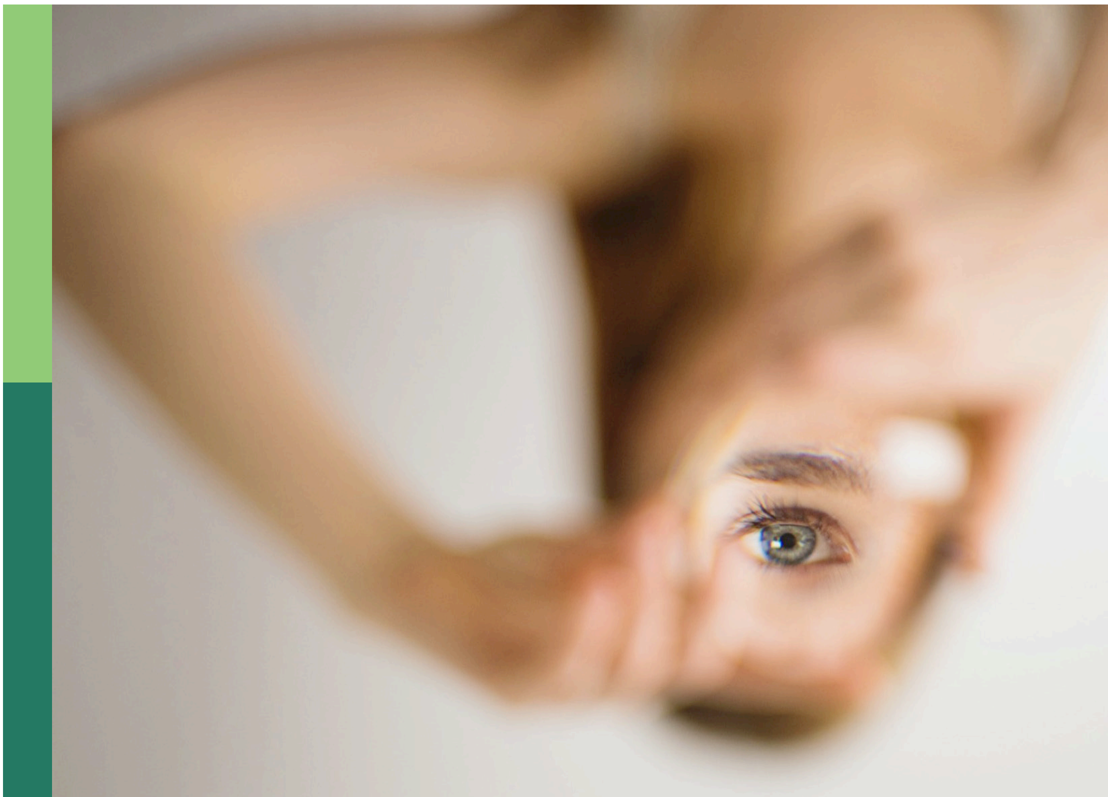
frontiers | Research Topics

From sub-lexical to discourse-level effects in bi- and multilingual language processing

Edited by
Katarzyna Dziubalska-Kolaczyk, Monika M. Polczynska and Katarzyna Jankowiak

Published in
Frontiers in Psychology



ISSN 1664-8714
ISBN 978-2-8325-6004-4
DOI 10.3389/978-2-8325-6004-4

About Frontiers

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

Frontiers journal series

The Frontiers journal series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the *Frontiers journal series* operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews. Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

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

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers editorial office: frontiersin.org/about/contact

From sub-lexical to discourse-level effects in bi- and multilingual language processing

Topic editors

Katarzyna Dziubalska-Kolaczyk — Adam Mickiewicz University, Poland

Monika M. Polczynska — University of California, Los Angeles, United States

Katarzyna Jankowiak — Adam Mickiewicz University, Poland

Citation

Dziubalska-Kolaczyk, K., Polczynska, M. M., Jankowiak, K., eds. (2025). *From sub-lexical to discourse-level effects in bi- and multilingual language processing*. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-8325-6004-4

Table of contents

frontiers | Frontiers in Psychology

TYPE Editorial
PUBLISHED 03 February 2025
DOI 10.3389/fpsyg.2025.1558057

OPEN ACCESS

EDITED AND REVIEWED BY
Xiaolin Zhou,
Peking University, China

*CORRESPONDENCE
Katarzyna Jankowiak
✉ katarzyna.jankowiak@amu.edu.pl

RECEIVED 09 January 2025
ACCEPTED 21 January 2025
PUBLISHED 03 February 2025

CITATION
Jankowiak K, Połczyńska-Bletsos M and Dziubalska-Kołaczyk K (2025) Editorial: From sub-lexical to discourse-level effects in bi- and multilingual language processing. *Front. Psychol.* 16:1558057. doi: 10.3389/fpsyg.2025.1558057

Editorial: From sub-lexical to discourse-level effects in bi- and multilingual language processing

Katarzyna Jankowiak[1]*, Monika Połczyńska-Bletsos[2] and Katarzyna Dziubalska-Kołaczyk[1]

[1]Faculty of English, Adam Mickiewicz University, Poznań, Poland, [2]Department of Psychiatry and Biobehavioral Sciences, University of California, Los Angeles, Los Angeles, CA, United States

KEYWORDS

bilingualism, multilingualism, language processing, cognition, cross-language interaction

Editorial on the Research Topic
From sub-lexical to discourse-level effects in bi- and multilingual language processing

This Research Topic delves into the dynamics of bi- and multilingual language processing, emphasizing the diverse influences of multilingualism across sub-lexical, lexico-semantic, and discourse levels. As linguistic diversity expands globally, the psycholinguistic effects of bilingualism and multilingualism continue to gather scholarly attention. The present contributions not only highlight the complexity of cross-linguistic interactions, but also stress the necessity of an interdisciplinary approach, bridging psychology, sociology, and linguistics. Each study advances our understanding of how multilingual individuals process language, adapt cognitively, and manage linguistic resources, ultimately enriching the field's broader perspective on the varied and evolving nature of multilingual cognition.

Geng et al. examined the processing of English-derived Japanese loanwords among Chinese learners of Japanese. Their study focused on factors such as familiarity, phonological similarity, context, and English proficiency. Familiarity was found to significantly reduce cognitive load, enhancing recognition, while the effect of phonological similarity diminished with higher Japanese proficiency. This finding suggests that advanced learners increasingly access Japanese meanings directly, bypassing reliance on English cues. The results underscore the importance of considering L1–L2 interactions when developing effective multilingual vocabulary resources.

Kędzierska et al. explored vowel perception in multilingual speakers of Polish, English, and Norwegian using event-related potentials. They examined how the mismatch negativity (MMN) response varies between a speaker's native language (L1, Polish) and non-native languages (L2 and L3/Ln). Results revealed that L1 elicited a stronger MMN response compared to L2 (English) and L3 (Norwegian), suggesting that language status modulates early auditory processing. This study enriches our knowledge of multilingual phonological perception and the roles of proficiency, dominance, and age of acquisition on phonemic discrimination.

Kim and Nam investigated the neural mechanisms underlying foveal word recognition through interhemispheric inhibition, using Korean visual stimuli. Their findings support the Split Fovea Theory, demonstrating that divided hemispheric processing reduces cognitive redundancy and enhances word recognition efficiency. While the study focused on monolingual word processing, its insights into hemispheric coordination and inhibitory control are relevant to bilingualism. These findings deepen our understanding of bilingual language control, cognitive resource management, and neural adaptability in multilingual individuals.

Laure and Armon-Lotem examined how Hebrew L2 bilinguals process templatic words, revealing that L1 mechanisms influence L2 word processing. Through a Hebrew rhyme judgment task, both Hebrew-native and Hebrew-L2 adults (both Semitic and non-Semitic L1s) were studied. Results indicated that Hebrew-L2 speakers utilize L1 patterns but show L2-specific adaptations, particularly in phonological or morphological awareness depending on their L1 background. Altogether, the findings highlight the influence of cross-linguistic transfer on L2 processing and provide insights into the role of non-linear morphology in bilingual language processing.

Wang et al. examined the effects of study-abroad experience (SAE) on Chinese (L1)–English (L2) interpreting students' translation skills. The study found that SAE participants translated more quickly but with more errors, indicating a speed-accuracy trade-off. Additionally, SAE participants demonstrated balanced bidirectional translation abilities, while non-SAE participants showed a preference for translating from L2 to L1. These findings suggest that SAE enhances cognitive flexibility and language-switching efficiency, pointing to the importance of immersive environments in interpreter training.

Baron et al. investigated grammatical gender processing in Spanish monolingual and Spanish–English bilingual children using eye-tracking. Testing children aged 5–10, they examined the use of gender cues in a visual world paradigm with grammatical and ungrammatical article-noun pairings. Results showed that bilinguals with greater English exposure were slower and less accurate in using gender cues than their monolingual peers. The findings highlight the impact of cumulative English exposure on grammatical gender processing and language control in bilingual children.

Fan and Wang investigated how L2 learners process formulaic sequences (FSs) during writing tasks with differing topic familiarity. The study distinguished internal FSs, which learners retrieve as whole units, from externally assembled FSs. The findings showed that high-proficiency learners more frequently retrieved and modified internal FSs, especially on familiar topics, indicating syntactic flexibility. In contrast, lower-proficiency learners assembled FSs word-by-word. These results suggest that L2 instruction should focus on internalizing FSs and promoting syntactic adaptability to enhance learners' writing fluency and accuracy, tailored to proficiency and topic familiarity.

Kul examined how Polish learners perceive reduced English forms, focusing on the effects of lexical context, phonetic reduction type, and musical background. The author found that lexical context and phonetic density significantly enhanced perception accuracy and speed, while musical training offered limited benefit, slightly improving reaction times but not accuracy. These findings suggest that language instruction should emphasize listening exercises featuring context-rich, naturally reduced speech patterns rather than idealized textbook clarity, helping learners navigate and understand authentic, connected spoken language more effectively.

Finally, Malarski et al. explored dialect use and style-shifting in the speech of Polish migrants in Norway, focusing on the acquisition of Norwegian (L3). Through sociolinguistic interviews in Oslo and Tromsø, the authors examined how first-generation migrants develop sensitivity to local dialects. Findings revealed that speakers vary in their use of regional features, with some acquiring dialectal forms similar to native speakers while others display less dialect use. The study offers valuable insights into multilingual dialect acquisition and sociolinguistic variation.

The articles in this Research Topic highlight the rich and varied landscape of multilingual language processing, spanning phonology, morphology, syntax, and discourse. Collectively, they underline the role of cognitive and linguistic factors, including proficiency, cross-linguistic influence, and immersion, in shaping language processing. Findings reveal how language status affects phonemic processing, how immersive experiences refine translation skills, and how bilingualism influences grammatical gender sensitivity and formulaic sequence usage. Ultimately, this body of work illustrates the adaptive and dynamic mechanisms of multilingual cognition, providing a nuanced understanding of how multilingual individuals navigate and manage complex linguistic resources across various contexts.

Author contributions

KJ: Writing – original draft, Writing – review & editing. MP-B: Writing – original draft, Writing – review & editing. KD-K: Writing – review & editing.

Conflict of interest

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

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

TYPE Original Research
PUBLISHED 06 June 2023
DOI 10.3389/fpsyg.2023.1164510

OPEN ACCESS

EDITED BY
Katarzyna Jankowiak,
Adam Mickiewicz University, Poland

REVIEWED BY
Hamutal Kreiner,
Ruppin Academic Center, Israel
Ehab W. Hermena,
Zayed University, United Arab Emirates

*CORRESPONDENCE
Yael Laure
✉ yaellaureac@gmail.com

RECEIVED 12 February 2023
ACCEPTED 02 May 2023
PUBLISHED 06 June 2023

CITATION
Laure Y and Armon-Lotem S (2023) Hebrew-L2 speakers process auditory templatic words through their L1 processing mechanism with awareness of L2. *Front. Psychol.* 14:1164510.
doi: 10.3389/fpsyg.2023.1164510

Hebrew-L2 speakers process auditory templatic words through their L1 processing mechanism with awareness of L2

Yael Laure[1]* and Sharon Armon-Lotem[1,2]

[1]Department of English Literature and Linguistics, Bar-Ilan University, Ramat Gan, Israel, [2]The Multidisciplinary Brain Research Center, Bar-Ilan University, Ramat Gan, Israel

Bilingualism involves cross-linguistic influence (CLI) prompted by communicative function, which impacts the activation of the bilingual's L1/L2 language processing mechanisms. The current study examines the extent of CLI when semantic information is reduced. Semitic languages are known for their templatic words composed of intertwined sub-lexical root and template morphemes, entailing non-linear morphological processing. As the roots constitute the semantic core, comprehension was found to impact morphological processing among Hebrew-L2 readers. Herein, we assessed the processing mechanism activated among adult Hebrew-L2 bilinguals in an auditory rhyme judgment task that requires linear processing. The task was provided with Hebrew templatic word pairs comprising accentuated (meta)linguistic information irrespective of semantics: phonological co-occurrence restrictions (root), grammatical information of vocalic melodies (template), and contrastive stress. We hypothesized that CLI in Hebrew-L2 speakers would be reflected in low accuracy rates in rhyming pairs when linguistic information is accentuated, indicating distraction from the linear processing due to activation levels of the L2 processing mechanism caused by competing linguistic cues drawn on transferred linguistic information. We compared the performance of 58 adult Hebrew native speakers with 54 Hebrew-L2 speakers with Semitic and non-Semitic-L1. The findings demonstrate that Hebrew-L2 speakers performed the task using their L1 processing mechanism with varying activation levels of L2, showing awareness of the morphological processing due to the vocalic melody for non-Semitic-L1 and awareness of contrastive stress for Semitic-L1. The results confirm CLI also when semantics is reduced, elucidating how much CLI modulates the bilingual's language processing mechanism.

KEYWORDS

word processing mechanism, sub-lexical morphemes, cross-linguistic influence, metalinguistic awareness, Hebrew-L2, rhyme judgment task

1. Introduction

Cross-linguistic influence (CLI) in bilinguals includes transfer of linguistic information between L1 and L2 and activation of these languages' processing mechanisms. The current study draws on the language mode hypothesis (Grosjean, 2001) and the unified competition model (MacWhinney, 2005) in a complementary manner, to explore the cross-linguistic influence of grammatical information when semantic information is reduced. Semitic languages are known for their templatic words composed of intertwined sub-lexical root and template morphemes, entailing language-specific non-linear processing. Examining activation of language mode in templatic word decomposition in an auditory rhyme judgment task, we probe CLI concerning metalinguistic information of the root and template and contrastive stress among Hebrew-L2 adult speakers, whose L1 is either Semitic or non-Semitic, compared to Hebrew native speakers. The influence was assessed in two-resolution levels: general awareness of the morphological processing and particular awareness of the root and template morphemes and stress. We hypothesized that CLI would occur when the linguistic information taps into induced or learned awareness but not subliminal awareness. We also hypothesized that the results would be affected by the bilingual's L1; therefore, we analyzed the results accordingly. To address our aims, we first present the unified competition model and the language mode hypotheses. Next, we explain the Hebrew templatic words and the sub-lexical root and template morphemes, and finally, we present the design of this study.

Bilingualism involves L1-L2 interplay in the bilingual's processing mechanism. The unified competition model (MacWhinney, 2005) premises that languages are used in the service of communicative function. Transfer of linguistic information between L1 and L2 increases the form-function matching possibilities, resulting in conflicting linguistic cues that compete to be selected by the language processing mechanism. The winning cue is the stronger and more reliable one, due to entrenchment obtained by neural circuits formed by (co)activation of the specific linguistic information. However, activation of the bilingual's languages is also affected by non-linguistic parameters, such as whether the bilingual is being spoken or listened to, the bilingual's proficiency (socioeconomic status, the usual mode of interaction), the situation (physical location, degree of formality), the form or content of the message (the topic, visual/aural modality, vocabulary), the function of the language act (communicating vs. creating social distance, taking part in an experiment), and research factors (aim, task used, organization, and type of stimuli) (Grosjean, 1998, 2001). The language mode hypothesis (Grosjean, 2001) holds that the bilingual's languages are activated in changeable levels at a given point in time, ranging from no activation to full activation, with one of the languages being the governing processing mechanism, which is also changeable. Thus, L2 can be either free of L1 interference or filtered through L1, or L1 and L2 change each other.

Studies among bilinguals with Hebrew-L2 have shown transfer of functional linguistic information. Influence of L2 on L1 has been demonstrated among Russian-Hebrew and Dutch-Hebrew children who applied non-native-like processing strategies of syntactic cues: case vs. word order (Janssen et al., 2015). Influence of L1 on L2 has been shown in lexical retrieval mechanism of object relative clauses among Russian-Hebrew children (Botwinik et al., 2016). Bidirectional transfer has been found in semantic word processing by sharing translation among English-Hebrew speakers who learned Hebrew as a first or second language (Degani et al., 2011). Moreover, although transfer is predicted to be minimized when language-specific properties exist in one of the bilingual's languages but not in the other (MacWhinney, 2005), morpho-syntactic features such as definiteness (only Hebrew), syntactic aspect (only Russian), and accusative case (both Russian and Hebrew) have been shown in bidirectional L1-L2 influence among Russian-Hebrew bilingual children (Meir et al., 2017).

Unlike children, who acquire L2 more completely than adults, adults need to coactivate L1-L2 linguistic knowledge by utilizing (meta)linguistic awareness, which is achieved by attention to the similarities and differences between the languages (MacWhinney, 2013; Bley-Vroman, 2018). The current study addresses CLI of language-specific metalinguistic information (apart from meaning) concerning the root and template and non-linear processing of templatic words among Hebrew-L2 adult bilinguals whose L1 is either Semitic or non-Semitic. According to the unified competition model, transfer of information or competition are not expected to occur because (i) the communicative function is not involved in this study, (ii) transfer is minimized when the linguistic information exists only in one of the bilingual's languages, as is the case with the non-Semiticn-L1, and (iii) Semitic languages share this particularity (McCarthy, 1981), rendering competition redundant, as is the case with the Semitic-L1. However, given that non-linguistic parameters also play a role and L1-L2 activation is dynamic, activation of the L2 processing mechanism, as evidence of linguistic transfer, may be seen due to the experiment requirements, modality, stimulus type, and (meta)linguistic awareness level. Therefore, these parameters should be handled and scrutinized with precision.

Semitic languages are known for their templatic words composed of intertwined two sub-lexical morphemes. The root, 2–6 (most common 3) consonantal phonemes, submitted to phonological co-occurrence restrictions, provide the semantic core. The template, vocalic pattern with or without fixed consonants, provide functional and grammatical information. Each morpheme may have more than one meaning. The word's meaning is a result of the joined morphemes and context. For example, in Hebrew, the words *xaʃav* (he thought/accountant) and *xiʃev* (he calculated) are composed of the root *xʃv* (think/calculate) and the templates *-a-a-* and *-i-e-* (verbs in the past tense/*-a-a-* also denotes profession; historically, a different template), and *maxʃev* (computer) with the *ma—e-* template with fixed consonants (denoting a tool or a place). While the nominal system comprises about a hundred templates, the verbal system comprises seven templates of verbal structures called binyanim: in general, three for active voice (Pa'al, Pi'el, Hif'il), three for passive voice in mirror relations to the active ones (Nifa'l, Pu'al, Huf'al), and one mostly reflexive, reciprocal, and inchoative (Hitpa'el). The root-template derivational relations in the binyanim are relatively firm, though not without exceptions. An especially stringent active–passive derivational relation is the one of binyanim Pi'el-Pu'al, e.g., *xiʃev-xuʃav* (calculated-was calculated). Derivations in Pa'al-Pu'al that share the same root may be semantically related (aka opaque) like *xaʃav-xuʃav* (thought-was calculated) or non-related like in *pasal-pusal* (canceled-was sculptured); however, they do not hold active–passive relations of the same action.

Composing words by intertwining the root and template sub-lexical morphemes entails non-linear processing, a Semitic linguistic particularity (aka the root-template mechanism). This contrasts with the universal approach that words are concatenative strings of phonemes. For example, the word *good* in English is a concatenative string of the phonemes /g/+/ʊ/+/d/, and *good*<u>*ness*</u> is a concatenative composition of the stem word *good* and the suffix *-ness*. Linear vs. non-linear processing of templatic words provokes an ongoing linguistic debate. Linguistic theories that advocate a stem-base lexicon deny the independent status of the root and template morphemes, providing linear processing mechanisms for Semitic templatic words (Aronoff, 1976; Bat-El, 1994, 2003; Benmamoun, 2003; Heath, 2003), so do computational theories, such as the Optimality Theory (Ussishkin, 2000, 2003). However, psycholinguistic research brings evidence for the independent representation of the root and template and non-linear processing, required for access to the lexicon and reading. Using masked priming paradigms experiment with lexical decision, repetition, and recognition tasks with words and non-words with real roots and with invented roots conditions, visual morphological priming effects were found among Hebrew native speakers in the verbal system (Deutsch and Frost, 2003) and in both verbal and nominal domains (Yablonski and Ben-Shachar, 2016), as real roots in non-words resulted in lower accuracy and longer response time compared to non-words with non-real roots or real words. Priming effects were also found in cross-modality models (visual and auditory), showing semantic-dependent priming effects in Arabic (Marslen-Wilson et al., 1994) or semantic-increasing effects in Hebrew (Gafni et al., 2019). Priming effects were also shown in the transposed root condition (phonemes in a different order) in visual experiments comparing Hebrew with English (Velan and Frost, 2007) or comparing Hebrew words of Semitic vs. non-Semitic origin (Velan and Frost, 2011), as well as in auditory repetition tasks (Oganyan et al., 2019).

Although these studies indicate awareness of Hebrew native speakers of the root and template morphemes, this awareness is semantic-dependent, as the studies used meaning-prompted experiments. The use of the printed word in Hebrew accentuates the root, tapping into the semantic core, as Hebrew is an abjad language. In addition, the use of printed words in a sentence requires semantics since the output is context-dependent. Moreover, the three-condition assessment consisting of real words, non-words with real roots, and non-words with non-real/illegal roots, as well as the transposed roots, in naming and lexical decision tasks compel semantic involvement, as the outcome taps into vocabulary knowledge. The question is, does non-linear processing occur when semantics is reduced, that is, due to (meta)linguistic information of the sub-lexical morphemes?

The linguistic impact of the root and template on non-linear processing of templatic words has been examined among Hebrew native speakers in a study using phonological awareness rhyme judgment task (saying if a pair of words rhymes) with auditory Hebrew templatic word stimuli with accentuated roots and templates for their linguistic information regardless of the meaning (Laure and Armon-Lotem, 2023a). This metalinguistic awareness measure, where semantics was reduced, creates an arena where the sub-lexical morphemes root and template compete with the syllables and sub-syllabic units for the processing mode: non-linear vs. linear, respectively. Success (accuracy) in this task indicates linear processing since the task requires parsing words linearly to syllables and phoneme discrimination. Hence, the low accuracy shown in rhyming pairs points to distraction from the linear processing. Low accuracy was shown in rhyming pairs where roots were identical in a pair, enabling the vocalic melody templates to stand out for their function, and when roots were transposed in a pair, minimizing the phonological feature realization in codas. The authors found the vocalic melody templates to have an abstract representation tapping into metalinguistic awareness of lexical-syntactic information and that transposed roots accentuate the application of the roots' phonological co-occurrence restrictions (Greenberg, 1950), tapping into subliminal linguistic knowledge of the computational phonological system.

A follow-up study comparing Hebrew native speakers with non-Hebrew speakers corroborated different processing mechanisms for Hebrew and non-Semitic non-Hebrew speakers (Laure and Armon-Lotem, 2023b). The non-Hebrew speakers processed rhyming and non-rhyming CVCVC pairs equally according to sub-syllabic units and phoneme similarity hierarchy in the final syllable, i.e., accuracy in recognizing rhymes decreases with the increase in similar phonemes in the final syllable unless the final syllables are identical (Lenel and Cantor, 1981). No awareness was shown among non-Hebrew speakers of the phonological co-occurrence restrictions or the abstract representation of the vocalic melody. Notably, the non-native Hebrew speakers scored low in non-rhyming identical templatic word pairs with contrastive stress (e.g., ***be**rex*-*be**rex*** (knee-blessed); stressed syllable in bold), likely because the contrastive stress is difficult to be perceived by speakers of languages with non-contrastive stress, like French (Segal and Kishon-Rabin, 2019). These results indicate that the grammatical information of the vocalic melody and the phonological co-occurrence restrictions are part of the (un)conscious metalinguistic knowledge of the Hebrew speaker. Having demonstrated that the root and template impact non-linear processing regardless of semantics among Hebrew native speakers, the question is, does this linguistic L2 particularity transfer and impact Hebrew-L2 bilinguals?

The Hebrew sub-lexical morphological non-linear processing has been examined among Hebrew-L2 readers in reading experiments that included words and non-words manipulated by four combinations of different/similar roots and patterns, reflecting on cross-linguistic influence (Norman et al., 2016, 2017). A study involving Hebrew-L1 and proficient Hebrew-L2 readers has found that morphological processing preceded lexical access for both Hebrew-L1 and proficient Hebrew-L2 readers from Indo-European and Semitic-L1 backgrounds, evident by processing strategies tuned to the root and template morphological processing in reading tasks (Norman et al., 2017). By contrast, a study including Hebrew-L2 learners in the early stages of learning has shown that the participants were modulated by L1 morphological background: L1-Indo-European beginning learners demonstrated sensitivity to the word pattern and word edges but not to the roots, and L1-Semitic beginning learners showed sensitivity to the fact that the word is the ensemble of both morphemes, as in Arabic, but not additive sensitivity to the root or template (Norman et al., 2016). However, although the morphological processing stands out in these experiments, it is impossible to disconnect the cognitive requirement associated with the written word in Hebrew, where the roots are salient, as words are written without vowels (abjad). Therefore, the question is, does cross-linguistic influence occur concerning the linguistic information associated with Hebrew templatic words regardless of semantics among Hebrew-L2 speakers, whose L1 is either Semitic or non-Semitic?

The present study explores this question using an auditory rhyming judgment task provided with Hebrew templatic words and comparing the performance of Hebrew-L2 speakers with Hebrew native speakers. Specifically, we examine the strength of L2 (meta)linguistic information concerning the morphological processing in general and the sub-lexical morphemes root and template and contrastive stress in particular by assessing activation of language modes in decomposing different templatic word stimulus types. The rationale for this study relies on the idea that the processing mechanism used by the bilinguals to decompose templatic words can reflect on the competition between conflicting linguistic information of sub-lexical units, i.e., root and template morphemes vs. syllables and sub-syllabic units. The auditory rhyme judgment task is utilized to designate the language modes' activation state (governing mechanism) and levels, while the stimulus types specify the sort of the transferred linguistic information of the root and template and stress.

Rhyme judgment tasks are part of a battery of tests that assess phonological awareness—the ability to understand that words are a series of sounds apart from their meaning. Phonological awareness is a language-universal construct (Branum-Martin et al., 2015) and has been shown to contribute to L2 consolidation (Zion et al., 2019). Recognizing rhymes requires identifying identical final words' vowel and consonant phonemes, which is successfully performed when parsing the words linearly into syllables, sub-syllabic units, and phonemes (Lewkowicz, 1980). Thus, the task compels linear processing, even more so when the stimuli are auditory since phonemes in spoken words are heard sequentially. Utilizing this task with templatic word stimuli creates an environment for competition between syllabic and morphemic sub-lexical units. The competition is even more difficult since the task requires mechanical decoding irrespective of semantics, while the roots get linguistic strength from the meaning. Moreover, non-linear processing is not beneficial in this task, as disentangling two intertwined sub-lexical morphemes in each word in a pair and then discriminating and comparing phonemes in two different sub-lexical units is cumbersome, costlier, and prone to errors. The strength of the syllabic units is granted not only by the task requirements but also by the universal phonological awareness construct. In addition, linear processing is also easier than non-linear processing (Upasana et al., 2022) and entrenched, as ceiling effects in this task have been seen by age six (Fox and Routh, 1975; Lewkowicz, 1980).

Since morphological processing occurs before lexical access (Norman et al., 2017), we hypothesized that despite the non-communicative function, the sub-lexical units (syllables/sub-syllabic units vs. morphemes) would compete due to the mechanical processing function required by the experiment. Success in this task, measured in accuracy, is indicative of linear processing. Low accuracy in non-rhyming pairs may occur due to inaccurate phoneme discrimination or to rhyme perception, i.e., phonemes are discerned but not always considered rhyme breakers. However, low accuracy in rhyming pairs is not expected in adults. Therefore, we considered low accuracy in this experiment as a distraction from linear processing. Given the balance of power of the two kinds of sub-lexical units, low accuracy in rhyming templatic word pairs would manifest cross-linguistic influence, indicating the strength of the linguistic L2 language-specific particularity of the sub-lexical root and template morphemes. Comparing the performance of Hebrew-L2 with Hebrew native speakers, we used two-resolution levels to assess cross-linguistic influence. The first aimed to capture a general awareness of the morphological processing of L2 Hebrew templatic words by examining accuracy in rhyming vs. non-rhyming pairs between and within stimulus types. The second focused on the linguistic information examined of each sub-lexical morpheme: phonological co-occurrence restrictions in transposed roots and the grammatical function of the vocalic melody in the template, both in rhyming pairs, and the contrastive stress in non-rhyming, non-stress-matched pairs.

Stimuli were sorted with linguistic precision to tap into the functional linguistic information irrespective of semantics. To avoid balanced-out results, we focused on one template type. All pairs were structure-matched with real roots. Manipulations emphasizing the root, template, and stress were achieved by combinations of different/identical real templates and different/identical real roots [+/-root,+/-template] in pairs, including transposed roots, vocalic melody templates, some of which violate binyanim relations, and non-stress-matched pairs. CLI was expected based on the (meta)awareness level, i.e., subliminal or induced or conscious. Specifically, we mainly expected salience of L2 linguistic information concerning the vocalic melody templates since the derivational function of vocalic melody is known, albeit to a lesser degree, also in non-Semitic languages (e.g., in English: choose-chose, begin-began-begun, etc.). In addition, the function of the template is prominent in the verbal system and taught, albeit on a semantic basis, in formal education, including Hebrew-L2 schools. Furthermore, it can be induced by usage. By contrast, awareness of the phonological co-occurrence restrictions was not predicted, as they are part of the subliminal computational system (regardless of whether submitted to grammar or statistical learning) (Berent, 2017). Awareness of the contrastive stress was partially expected. Unlike phonological awareness, stress is a language-specific construct (Branum-Martin et al., 2015). Acquisition of a second language facilitates awareness of stress when stress distinguishes between word meanings (Segal and Kishon-Rabin, 2019). However, the meaning of the words does not play a role in this task, which might impact the activation of L2. In addition, unlike Spanish, for example, where the contrastive stress is seen in the written modality, contrastive stress in Hebrew appears only in the oral/aural modality (it has no realization in Hebrew in the visual and written modality), and it can be resolved in context. Thus, the opportunities to induce or learn this awareness are reduced and modality-dependent.

Generally, we expected that cross-linguistic influence would be seen in the Hebrew-L2 speakers not just by the levels of (meta)linguistic awareness pronounced in the two-resolution levels but also by the level of similarity/difference between L1 and L2. Forces based on L1-L2 similar/different mechanisms and linguistic information are not equal when the Hebrew-L2 bilinguals have Semitic-L1 or non-Semitic-L1. It is difficult to tell which language mechanism is activated when L1 is Semitic, due to shared characteristics, and when L1 is non-Semitic, as linguistic transfer is expected to be minimized when the linguistic information pertains to only one of the bilingual's languages

(MacWhinney, 2005). Therefore, precision is required. Although templatic words, root and template, and non-linear processing are common in Semitic languages (McCarthy, 1981), they have different manifestations. For example, the active–passive vocalic melody templates differ in length: The vocalic melody in the verbal system in Arabic includes three vowels [*-a-a-a* (active) -*u-i-a* (passive)], as opposed to two vowels [*-i-e-* (active) *-u-a-* (passive)] in Hebrew. In addition, beginner Hebrew-L2 learners with Semitic-L1 did not show sensitivity to the root or template in reading (Norman et al., 2016). In addition, as mentioned above, functional vocalic melody pattern is not exclusive to Semitic languages, although much more pervasive. Also, contrastive stress is language-specific, regardless of language family affiliation: contrastive stress in Hebrew, Spanish, and English vs. non-contrastive in French and Arabic (Segal and Kishon-Rabin, 2019). Therefore, we addressed the results using a multi-layer analysis: one of the Hebrew-L2 as a whole and the other separated by the bilinguals' L1.

Altogether, the task chosen, the precision in stimulus types, the two-resolution evaluation, and the multi-layer analysis enable the examination of cross-linguistic influence, including language activation and transfer of (meta)linguistic information.

2. Materials and methods

2.1. Participants

A total of 186 adults participated online; all declared not to have hearing problems. Participants who responded to <85% of the stimuli were removed from the sample. 112 participants finished the experiment. 58 were Hebrew native speakers (Heb1) (ages 20–82 years, 39 female speakers), and 54 were Hebrew-L2 speakers (Heb2) (ages 21–82, 38 female speakers), with L1 including a Semitic language (Arabic) and non-Semitic languages. The participants filled in a questionnaire regarding demographic details, education in categories matching worldwide distinctions (some school, high school diploma, some college, undergraduate, graduate, and postgraduate), linguistic background information about the age of acquiring Hebrew, the number of years they use it, and their level of Hebrew in speech, reading, and writing on a 0–10 self-rating scale (Table 1). Speech level was important due to the auditory modality in this experiment. Since the experiment was performed online, the Hebrew native speakers were asked to self-rate their Hebrew level on a 0–10 scale in speech, reading, and writing, to verify that participants who defined Hebrew as their mother tongue could be defined as L1 Hebrew speakers (Table 1). The study was approved by the ethics committee of the Faculty of Humanities, Bar-Ilan University.

2.2. Stimuli

The experiment included 205 stimulus pairs, comprising 64 rhyming (R) and 141 non-rhyming (NR) auditorily structure-matched pairs. All words (some archaic) were examined to be valid using the Even-Shoshan dictionary (Even-Shoshan, 1974) and The Academy of the Hebrew Language site (https://hebrew-academy.org.il/. Accessed October 15, 2022). All bi-syllabic words are affix-free templatic words of full 3-consonantal roots assessed by the phonemic representation, without weak roots (where one of the root consonants is missing/not transparent) or geminate roots (e.g., tss, grr, etc.). We excluded the phonemes /ʔ/ (א), /ʕ/ (ע), /h/ (ה) since they may alter the auditory syllable structure (e.g., CVCV בנה (*bana*) or VCVC ענד (*anad*) or CV.VC דאג (*da'ag*) instead of CVCVC). We also excluded words with suffix-like final -VCs, such as /*-im*/, /*-ot*/ (plural morphemes), /*-on*/, /*-it*/ (diminutive), and /*-an*/ (personality characteristics/profession in Hebrew, and Accusative case in Arabic). Frequency of the words in the language was not considered as the experiment is metalinguistic awareness oriented. No impact of frequency was predicted for word decomposition as the target is technical parsing to isolate and compare the rime (-VC) and phonemes of the words in a pair.

The stimuli encompass mono-syllabic CVC pairs, used as control, and bi-syllabic structure-matched pairs of templatic words composed of roots and templates, including pairs with templatic fixed consonants mVCCVC and CVCVC pairs of [+/–CR,+/–VM] [CR for consonantal root; VM for vocalic melody (template)] combinations in pairs, including transposed roots, templates violating binyanim relations, and non-stress-matched pairs. The purpose of each stimuli type is detailed below (see Table 2 for examples).

1. CVC (62) pairs (15R/47NR), representing all identical and contrasting coda possibilities. We use them to ensure phoneme discrimination and the ability to recognize rhymes. This group is also a basis for comparison of R/NR processing in mono-syllabic vs. bi-syllabic pairs. These pairs are not templatic words.
2. mVCCVC (21) pairs (7R/14NR), nominal templates with initial consonants indicating tools or place. This type is used for comparison of R/NR processing with the CVCVC template, to pinpoint similarity or differences between two types of templatic words. Similarity between this type and the Baseline in CVCVC enables emphasizing the accentuated roles of the root and the template in the CVCVC pairs.

Six combinations of the 122 CVCVC pairs.

3. *Baseline[–CR,–VM] (B(–))*. 24 pairs (10R/14NR) with different CRs (ranging 0–2 out of 3) and VMs (ranging 0–1 out of 2) within the pair's words; used for setting the baseline for CVCVC pairs for representing phoneme variety and comparison of R/NR processing.
4. *Baseline[–CR,+VM] (B(+))*. 14 pairs (8R/6NR) with different CRs (ranging 0–2 out of 3) and identical VMs within the pair's words; used for comparisons of R/NR processing and for comparison with the transposed root pairs to establish the phonological impact of the transposed phonemes.
5. *Transposed-CR[–CR,+VM] (TCR)*. 40 pairs (8R/32NR) with roots sharing the same phonemes in different positions (transposed roots) and identical VMs. The transposed root stimulus type highlights the phonological co-occurrence restrictions in the roots. The non-rhyming pairs are four times

TABLE 1 Participants' background characteristics—means (SD) by language group.

	Hebrew native		Semitic-L1 Hebrew-L2*		Non-Semitic-L1 Hebrew-L2**		Hebrew-L2	
Categorical variables	*n*	%	*n*	%	*n*	%	*n*	%
Number and Gender	58		16		38		54	
Female	39	67.2%	15	93.8%	23	60.5%	38	70.4%
Male	19	32.8%	1	6.35%	14	36.8%	15	27.8%
Non-binary					1	2.6%	1	1.9%
Education								
Non-Academic	11	19%	3	18.8%	3	7.9%	6	11.1%
Academic	27	46.5%	13	81.2%	29	76.3%	42	77.8%
Non-specific	20	34.5%	–	–	6	15.8%	6	11.1%
Continuous variables	*Mean*	*(SD)*	*Mean*	*(SD)*	*Mean*	*(SD)*	*Mean*	*(SD)*
Age in years	41.83 (20–82)	*(14.61)*	26.19 (21–32)	*(3.33)*	49.66 (23–82)	*(19.48)*	42.70 (21–82)	*(18.94)*
Acquired Hebrew at the age of	N/A		6.81)	*(3.67)*	17.71	*(10.72)*	14.48	*(10.45)*
Number of years Hebrew is used	N/A		11.19	*(8.53)*	17.55	*(15.53)*	15.67	*(14.06)*
Hebrew speaking level	9.78	*(0.53)*	7.13	*(2.03)*	6.05	*(2.60)*	6.37	*(2.48)*
Hebrew reading level	9.79	(0.67)	7.56	*(2.07)*	5.95	*(3.07)*	6.43	*(2.89)*
Hebrew writing level	9.69	(0.68)	6.94	*(2.52)*	5.32	*(3.14)*	5.80	*3.04)*

*Arabic.

**By alphabetic order: Berber, Dutch, English, Farsi, French, German, Hungarian, Russian, and Spanish.

more the number the rhyme pairs since 3-consonantal roots have five swaps (named after the second word's alternation), four of which are non-rhyming pairs and one rhyming pairs (see Table 2 for swap examples). This stimulus type is used for comparison of R/NR processing with the Baseline(+) to evaluate the susceptibility of transposed roots to phonological restrictions compared to varying consonantal roots.

6. *Highlighted-VM[+CR,−VM] (HVM)*. 32 pairs (20R/ 12NR) with identical CRs and different VMs; used for comparison of R/NR processing, and also to examine in rhyming pairs the effect of binyanim relations: Pa'al-Pi'el, Pa'al-Pu'al, and MIX (no binyanim relations), as well as the impact of semantics by comparing pairs with vs. without semantic relatedness between the words in a pair (see Table 2 for examples).
7. *Stress[+CR,+VM] (Stress)*. 6 identical CRs and VMs in non-stress-matched (trochaic (in bold) vs. iambic) pairs, therefore non-rhyming pairs; used to examine awareness of the contrastive stress and its impact on processing in this task.
8. *Stress[+CR,+VM] (Stress(−))*. 6 identical CRs and different VMs in non-stress-matched (trochaic (in bold) vs. iambic) pairs, therefore non-rhyming pairs; used to examine the extent of the impact and awareness of the stress vis-à-vis the template.

The stimuli were recorded using the Audacity software in a feminine voice in a professional studio or a quiet room. All pairs started after 55 ms with 300 ms gap between the words in each pair. Following pre-trial pilot feedback, the pairs slightly varied in volume to keep participants alerted and focused on the task, and response time was limited to 2 seconds to avoid an unintuitive decision. The pairs were randomly divided into five sections, containing all stimulus types, and fully randomized within sections. The randomized order was similar for all participants.

2.3. Procedure

We spread the online experiment with information about the research's aim, requirements, and instructions available in eleven languages (by alphabetic order: Arabic, Chinese, English, Filipino, French, German, Hebrew, Italian, Portuguese, Spanish, and Russian) through emails and social media. The experiment was limited to computers only to increase uniformity in testing conditions. We added a hearing test to verify that the speakers of the participants' computers work. Participants filled in a questionnaire relating to demographic information and linguistic background, read the instructions about the task in which they were asked to follow their intuition, and performed a practice trial (as many times as they wanted) to familiarize themselves with the procedure and technical aspects of the real trial; no feedback was given to the participants in order not to impact or interfere in their rhyme judgment. Then, they started the experiment. The pairs were played sequentially, after a response was issued or 2 secs passed (displayed on a diminishing bar). The question "Does it rhyme?" and the Yes and No buttons constantly appeared on the screen in each section. Between sections, the participants watched a silent 15-sec nature video, which differed between sections but appeared in a fixed order.

TABLE 2 Examples of stimulus types.

		R		NR	
		Example	Translation	Example	Translation
1	CVC	*dag-xag*	fish-holiday	*sal-saʃ*	basket-threshold
2	mVCCVC	*mavreg-mazleg*	screwdriver-fork	*mivcar-migdal*	fortress-tower
3	Baseline[−CR,−VM](B(−))	*ʃimer-ʃalax*	preserved-sent	*ʃiger-sagar*	launched-closed
4	Baseline[−CR,+VM](B(+))	*mazal-kaval*	luck-complained	*gamad-ʔamat*	dwarf-dropped
5	TCR[−CR,+VM]				
	C132	–	–	*xatar-xarat*	rowed-engraved
	C321	–	–	*karas-sakar*	collapsed-surveyed
	C231	–	–	*zaram-ramaz*	flowed-hinted
	C312	–	–	*ʃitek-kiʃet*	paralyzed-decorated
	C213	*kalax-lakax*	flowed-took	–	–
6	HVM[+CR, −VM]	*dabur-dibur*	hornet-speech	*natav-nituv*	router-routing
	Binyanim Relations				
	MIX	*kaʃer-koʃer*	kosher-ties(verb)	–	–
	Pa'al -Pi'el	*lomed-limed*	learns-taught	–	–
	Pa'al-Pu'al	*saxak-suxak*	laughed-was played	–	–
	Semantic relatedness				
	Non-related	*bocer-bicer*	picked grapes-fortified	–	–
	Related	*ʃamen-ʃimen*	fat (adj)-greased	–	–
7	Stress[+CR,+VM]	–	–	*corex-**corex***	consumes-need
8	Stress[+CR, −VM]	–	–	*dover-**dever***	spokesman-plague

Trochaic stress in bold, otherwise iambic.
CR, consonantal root; HVM, highlighted-vocalic melody; NR, non-rhyming pairs; R, rhyming pairs; TCR, transposed-consonantal root; VM, vocalic melody.

Moving from one section to another required pressing "continue" and "start" buttons; thus, the pause length between sections was the participant's choice. The participants' answers were recorded in the database in their raw values: Yes, No. Then, the answers were converted to Correct (1) or Error (0) according to the following criterion: If both words of a pair have stress-matched identical final syllable's vowel and coda (–VC), the pair is a rhyming pair; otherwise, it is non-rhyming.

2.4. Statistical analysis

Since it was an online experiment of binary answers, to rule out malicious participants (pressing Yes/No blindly), we calculated the probability of each participant blindly answering the experiment question "Does it rhyme?." A participant that blindly chooses Yes (y) or No (n) has an empirical probability (p) p_y and p_n, respectively. Similarly, the probability of a pair in question (q) being a rhyme is $p_n \cdot q_n + p_y \cdot q_y = p$. We have m questions, and let the number of answers the participant answered correctly be k. With these parameters, we calculated the probability (or likelihood) of said person to answer k correct "random guesses" out of m questions using the binomial distribution formula: $\binom{m}{k} p^k (1-p)^{m-k}$. The results indicate that each subject has a probability of <0.05 (range from 0.051329145 to 2.7312E-48) of achieving their accuracy (see Table A1 in the Supplementary Material for more details). Hence, none of the results nearly 50% indicate by chance accuracy.

Correlations between the experiment's five sections (r range 0.716–0.926) indicate high stability and consistency; therefore, we analyzed the results without separating sections. We used the multilevel modeling (MLM) for repeated measures designs as it allowed us flexibility in modeling a more appropriate variance-covariance matrix, relative to the repeated measure ANOVA, and handling missing data using the full information maximum likelihood, performed using SPSS IBM V.27. In case of significant main effects or interaction effects, a further set of *post hoc* comparisons were performed. To avoid alpha inflation, a Bonferroni adjustment was applied.

3. Results

The results were analyzed *via* stimulus multi-layer analyses from two perspectives: comparison of Hebrew native speakers (Heb1) with the entire sample of Hebrew-L2 speakers (Heb2) and comparison of Hebrew native speakers with Hebrew-L2 speakers discerned by the participants' mother tongue, Semitic (S-Heb2) vs. non-Semitic (nS-Heb2), to explore the source of the cross-linguistic influence.

3.1. Processing rhyming vs. non-rhyming pairs

3.1.1. Mono-syllabic CVC vs. Bi-syllabic CVCVC pairs

To examine whether Heb1 and Heb2 process similarly rhyming and non-rhyming pairs in mono-syllabic CVC and bi-syllabic CVCVC stimuli, we compared accuracy according to length by rhyme value (CVC-NR/CVC-R/CVCVC-NR/CVCVC-R) × language group (Heb1/Heb2). The analysis revealed a significant effect of length by rhyme value [$F_{(3,193)} = 83.95$, $p < 0.001$] but not of language group [$F_{(1,336)} = 1.24$, $p = 0.266$] and an interaction effect [$F_{(3,193)} = 8.31$, $p < 0.001$]. *Post hoc* analysis indicated that in mono-syllabic pairs, Heb1 processed R and NR similarly ($p = 0.374$), whereas Heb2 scored significantly higher ($p < 0.001$) in R than NR. Heb1 scored significantly higher ($p = 0.014$) than Heb2 in NR but not in R ($p = 0.301$). In CVCVC, both Heb1 and Heb2 scored significantly lower in R than NR: Heb1 ($p < 0.001$), Heb2 ($p < 0.001$), with Heb2 scoring significantly higher ($p = 0.018$) in R and significantly lower ($p < 0.001$) in NR than Heb1 (Table 3).

The same analysis discerning Heb2 by the participants' L1 (language group (Heb1/nS-Heb2/S-Heb2) × length by rhyme value [CVC-NR/CVC-R/CVCVC-NR/CVCVC-R)] revealed significant effects of length by rhyme value [$F_{(3,187)} = 70.41$, $p < 0.001$], language group [$F_{(2,349)} = 17.81$, $p < 0.001$], and an interaction effect [$F_{(6,187)} = 7.56$, $p < 0.001$]. *Post hoc* analysis indicated that Heb1 scored similarly ($p = 0.336$) in R and NR in CVC but significantly lower ($p < 0.001$) in R than NR in CVCVC. In CVC, S-Heb2 scored significantly lower ($p < 0.001$) in NR than R, while nS-Heb2 approached significance ($p = 0.053$), with higher accuracy in R. No significant differences were shown in R in CVC between Heb1 and nS-Heb2 ($p = 0.218$), Heb1 and S-Heb2 ($p = 0.864$), and nS-Heb2 and S-Heb2 ($p = 0.483$). In CVCVC, both nS-Heb2 ($p < 0.001$) and S-Heb2 ($p < 0.001$) scored significantly lower in R than NR. Interestingly, nS-Heb2 scored significantly higher than S-Heb2 ($p = 0.006$) and Heb1 ($p < 0.001$), but no significant difference was shown between S-Heb2 and Heb1 ($p = 0.667$) (Table 3).

These results could be taken to show that Hebrew-L2 process mono- and bi-syllabic pairs similarly to Hebrew native speakers. However, when discerned by L1, non-Semitic-L1 is similar to Hebrew native speakers in CVC pairs, whereas Semitic-L1 is similar to Hebrew native speakers in CVCVC pairs.

3.1.2. Within the bi-syllabic stimulus types

Since putting all templates in one basket might conceal differences of specific particularities, we sought to examine whether the accuracy rates of Heb2 are similar to those of Heb1 in R and NR within and between the different stimulus types. Table 3 presents a comparison of language group (Heb1/Heb2) by stimulus type by rhyme value (CVC-NR/CVC-R/mVCCVC-NR/mVCCVC-R/TCR-NR/TCR-R/HVM-NR/HVM-R/B(+)-NR/B(+)-R/B(–)-NR/B(–)-R). The analysis revealed a significant effect of stimulus type by rhyme value [$F_{(11,204)} = 59.30$, $p < 0.001$] but not of language group [$F_{(1,986)} = 1.77$, $p = 0.184$] and an interaction effect [$F_{(11,204)} = 4.82$, $p < 0.001$]. *Post hoc* analysis showed that Heb1 scored significantly lower in R than NR in all the bi-syllabic pairs [mVCCVC ($p = 0.004$), TCR ($p < 0.001$), HVM ($p < 0.001$), B(+) ($p = 0.002$), and B(–) ($p < 0.001$)] but not ($p = 0.374$) in the mono-syllabic CVC. Heb2 scored significantly lower in R than NR in the bi-syllabic CVCVC types: TCR ($p < 0.001$), HVM ($p < 0.001$), and B(–) ($p < 0.001$), but similarly in R and NR in stimuli with identical templates in a pair: B(+) ($p = 0.327$) and mVCCVC ($p = 0.417$), and significantly higher ($p < 0.001$) in R than NR in CVC.

The same analysis discerning Heb2 by the participants' L1 (language group (Heb1/nS-Heb2/S-Heb2) × stimulus type [CVC-NR/CVC-R/mVCCVC-NR/mVCCVC-R/TCR-NR/TCR-R/HVM-NR/HVM-R/B(+)-NR/B(+)-R/B(–)-NR/B(–)-R)] revealed a significant effect of stimulus type by rhyme value [$F_{(11,197)} = 49.15$, $p < 0.001$], language group [$F_{(2,969)} = 22.61$, $p < 0.001$], and an interaction effect [$F_{(22,197)} = 5.93$, $p < 0.001$]. *Post hoc* analysis showed that Heb1 scored significantly higher in NR than R in all the bi-syllabic pairs (mVCCVC ($p = 0.003$), TCR ($p < 0.001$), HVM ($p < 0.001$), B(+) ($p = 0.002$), and B(–) [$p < 0.001$)] but not in the mono-syllabic CVC pairs ($p = 0.336$). S-Heb2 showed significant differences in all the bi-syllabic types; however, unlike the Heb1, not always the NR was higher than R: S-Heb2 scored significantly higher in R than NR in CVC ($p < 0.001$), mVCCVC ($p = 0.002$), and B(+) ($p = 0.001$) but significantly lower in R than NR in TCR ($p = 0.015$), HVM ($p < 0.001$), and B(–) ($p < 0.001$). nS-Heb2 also scored significantly lower in R than NR in the TCR ($p < 0.001$), HVM ($p < 0.001$), and B(–) ($p < 0.001$), but no significant differences were shown in mVCCVC ($p = 0.317$) and B(+) ($p = 0.284$), and in CVC, approaching significance ($p = 0.053$) with higher scores in R (Table 3).

These findings indicate that in the three stimulus types, namely TCR, HVM, and B(–), all the participants demonstrated a low accuracy rate in rhyming pairs. In the other three types [CVC, mVCCVC, B(+)], Semitic-L1 showed a low accuracy rate in non-rhyming pairs, while non-Semitic-L1 showed no difference between rhyming and non-rhyming pairs; both Semitic-L1 and non-Semitic-L1 contrast with Hebrew native speakers, whose scores were significantly lower in rhyming pairs.

3.2. Rhyming pairs in bi-syllabic stimulus types

3.2.1. Within language groups

Next, we sought to examine the impact each bi-syllabic stimulus type had on each language group. Based on the analysis in the previous section, we compared accuracy in bi-syllabic rhyming pairs between the five stimulus types [mVCCVC, B(+), B(–), TCR, HVM] in each language group (Table 3, letters). For Heb1, accuracy in mVCCVC was significantly higher than TCR ($p < 0.001$) and HVM ($p < 0.001$) but not than B(+) ($p = 0.855$) and B(–) ($p = 0.060$) although approaching significance; significantly lower in B(–) than B(+) ($p = 0.015$); significantly lower in HVM than TCR ($p = 0.047$), with both TCR and HVM significantly lower than B(–) [TCR ($p =$

TABLE 3 Means and (SD) of accuracy of rhyming and non-rhyming pairs of all stimulus types by language groups Heb1 vs. Heb2, with Heb2 discerned by L1.

	Heb1		Semitic-L1-Heb2		Non-Semitic-L1-Heb2		Heb2	
	R	NR	R	NR	R	NR	R	NR
CVC	0.91 (0.18)	0.87 (0.23)	0.91^{***} (0.13)	0.49 (0.32)	0.94 (0.08)	0.86 (0.22)	0.94^{***} (0.09)	0.75 (0.30)
Total CVCVC	0.46^{***} (0.24)	0.91 (0.10)	0.43^{***} (0.16)	0.76 (0.14)	0.63^{***} (0.24)	0.84 (0.16)	0.57^{***} (0.24)	0.82 (0.16)
Bi-syllabic pairs								
mVCCVC	0.72^{*}_{a} (0.29)	0.86 (0.21)	0.77^{*}_{a} (0.22)	0.50 (0.26)	0.74_{a} (0.30)	0.80 (0.18)	0.75_{a} (0.28)	0.71 (0.25)
CVCVC types								
Baseline[–CR,+VM]	0.73^{*}_{a} (0.19)	0.85 (0.23)	0.78^{**}_{a} (0.12)	0.53 (0.25)	0.75_{a} (0.18)	0.80 (0.21)	0.76_{a} (0.16)	0.72 (0.25)
Baseline[–CR,–VM]	$0.62^{***}_{a,b}$ (0.25)	0.98 (0.04)	$0.59^{***}_{a,b}$ (0.32)	0.92 (0.11)	$0.68^{***}_{a,c}$ (0.30)	0.93 (0.12)	$0.65^{***}_{a,b}$ (0.30)	0.92 (0.12)
TCR	0.44^{***}_{c} (0.33)	0.92 (0.14)	0.50^{*}_{b} (0.28)	0.71 (0.26)	$0.66^{***}_{a,b}$ (0.28)	0.92 (0.12)	0.61^{***}_{b} (0.29)	0.86 (0.20)
HVM	0.32^{***}_{c} (0.35)	0.95 (0.09)	0.22^{***}_{c} (0.24)	0.90 (0.15)	$0.55^{***}_{b,c}$ (0.38)	0.81 (0.32)	0.45^{***}_{c} (0.38)	0.83 (0.28)

Significance in comparisons of R with NR in the same language group and stimulus types are marked by asterisks according to conventional critical *P*-values: $p < 0.05$, $p < 0.01$, and $p < 0.001$. Means in the same column that do not share the sub-script are significantly different.
Heb1, Hebrew native speakers; Heb2, Hebrew-L2 speakers; HVM, highlighted-vocalic melody; NR, non-rhyming pairs; R, rhyming pairs; TCR, transposed-consonantal root.

0.001), HVM ($p < 0.001$)] and B(+) [HVM and TCR ($p < 0.001$)]. For the Heb2, accuracy in mVCCVC was significantly higher than TCR ($p = 0.017$) and HVM ($p < 0.001$) but not than B(+) ($p = 0.910$) and B(–) ($p = 0.072$) although approaching significance, significantly lower in B(–) than B(+) ($p = 0.024$); significantly lower in HVM than TCR ($p = 0.013$), B(+) ($p < 0.001$) and B(–) ($p = 0.001$), and significantly lower in TCR than B(+) ($p < 0.001$) but not B(–) ($p = 0.490$). Breaking down the Hebrew-L2 by the participants L1 showed that the similar trends between Heb1 and Heb2 were due to the Semitic-L1 Hebrew-L2 speakers. S-Heb2 demonstrated a similar trend to Heb1, with accuracy in mVCCVC significantly higher than TCR ($p < 0.010$) and HVM ($p < 0.001$) but not B(+) ($p = 0.948$) and B(–) ($p = 0.073$) although approaching significance; significantly lower in B(–) than B(+) ($p = 0.027$); significantly lower in HVM than B(+) ($p < 0.001$), B(–) ($p = 0.001$), and TCR ($p = 0.016$); and TCR lower than B(+) ($p = 0.002$), but not B(–) ($p = 0.369$). In contrast, nS-Heb2 demonstrated accuracy significant lower in HVM than mVCCVC ($p = 0.009$) and B(+) ($p = 0.002$), without any other differences [mVCCVC–B(+) ($p = 0.927$), mVCCVC–B(–) ($p = 0.327$), mVCCVC–TCR ($p = 0.240$), TCR–HVM ($p = 0.139$), TCR–B(+) ($p = 0.144$), TCR–B(–) ($p = 0.804$), HVM–B(–) ($p = 0.077$), and B(+)–B(–) ($p = 0.206$)].

These findings indicate that the Semitic native speakers (Hebrew native speakers and Semitic-L1) show a similar cascade of accuracy mVCCVC=B(+)>B(–)>(Heb1)/ = (S-Heb2)TCR>HVM, which differs from the results of the non-Semitic-L1 whose accuracy is lower only in HVM compared to mVCCVC and B(+). Two major differences between non-Semitic-L1 and Semitic native speakers are the differences shown between B(+) and B(–) and among Semitic native speakers between TCR and HVM, which are not shown among non-Semitic-L1 speakers (Figure 1, colored stars).

3.2.2. Between language groups

Next, we wanted to test whether the accuracy rate in rhyming pairs in the different stimulus types differed between language groups. Comparing the rhyming pairs from the previous analysis of the different bi-syllabic stimulus types between Heb1 and Heb2 indicated no difference between language groups in mVCCVC ($p = 0.543$), B(+) ($p = 0.385$), and B(–) ($p = 0.525$). However, significant differences were shown in the TCR ($p = 0.004$) and approaching significance ($p = 0.054$) in HVM (Figure 1). Breaking down the Heb2 to non-Semitic-L1 and Semitic-L1 speakers showed that this was due to the non-Semitic-L1-Heb2 only: In TCR, nS-Heb2 scored significantly higher than Heb1 ($p = 0.001$) but not than S-Heb2 ($p = 0.085$), without a difference between Heb1 and S-Heb2 ($p = 0.476$); and in HVM, nS-Heb2 scored significantly higher than Heb1 ($p = 0.002$) and S-Heb2 ($p = 0.002$), without a difference between Heb1 and S-Heb2 ($p = 0.321$) (Figure 1, asterisks in Data Table).

These findings indicate that both groups of Semitic native speakers (Heb1 and S-Heb2) scored lower than non-Semitic-L1 in the HVM. Interestingly, in TCR, only Hebrew native speakers scored lower than non-Semitic-L1, while Semitic-L1 showed no difference from either non-Semitic-L1 or Hebrew native speakers.

3.3. Varying vs. transposed-consonantal roots

Next, we wanted to probe if the lack of difference for nS-Heb2 between B(+) and TCR remains when the vocalic melody is identical in the entire stimuli sample and not only between pairs by removing the potential impact of grammatical information pronounced in the VM and thus better scrutinizing the impact of the phonological co-occurrence restrictions accentuated in the transposed pairs. To this end, we compared pairs sharing the VM *-a-a-*, half with varying consonantal roots (B(+)), and half with transposed-consonantal roots (TCR) in a pair. Table 4 presents a comparison of stimulus type (Transposed/Varying) × language group (Heb1/nS-Heb2/S-Heb2).

The analysis revealed a significant effect of stimulus type [$F_{(1,193)} = 30.27$, $p < 0.001$], language group [$F_{(2,193)} = 7.33$, $p = 0.001$], and an interaction effect [$F_{(1,193)} = 4.94$, $p = 0.008$]. *Post hoc* analysis showed significant differences between varying and transposed CRs

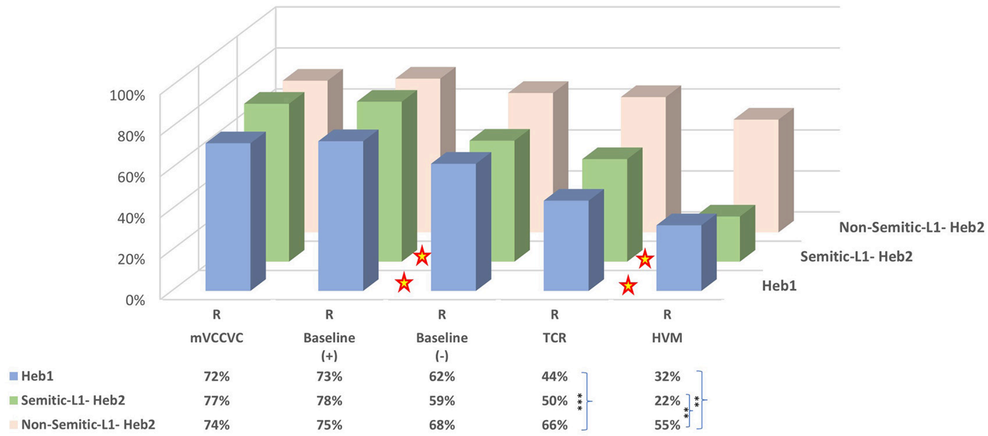


FIGURE 1
Rhyming bi-syllabic pairs by language groups (distinguished by L1). Colored stars indicate significant differences at $p < 0.05$ between B(−) and B(+) and between TCR and HVM for Heb1 and S-Heb2 language groups. Asterisks in the Data Table indicate significant differences between language groups marked by the curly brackets (in TCR and HVM) according to conventional critical P-values: $^*p < 0.05$, $^{**}p < 0.01$, and $^{***}p < 0.001$.

TABLE 4 Means and (SD) of varying vs. transposed-consonantal roots in *-a-a-* rhyming pairs.

Heb1		S-Heb2		nS-Heb2		Heb2	
Baseline	Transposed	Baseline	Transposed	Baseline	Transposed	Baseline	Transposed
0.76*** (0.23)	0.43_A (0.33)	0.78* (0.13)	0.51_a (0.32)	0.78 (0.22)	0.70_b (0.28)	0.78*** (0.19)	0.65_B (0.31)

Significance in comparisons of varying with transposed-consonantal roots in the same language group are marked with asterisks according to conventional critical P-values: $p < 0.05$, $p < 0.01$, and $p < 0.001$.
Means in the same row that do not share the sub-script are significantly different.
Lowercase indicates differences between Heb1, non-Semitic-L1-Heb2, and Semitic-L1-Heb2.
Uppercase indicates differences between Heb1 and Heb2.
Heb1, Hebrew native speakers; Heb2, Hebrew-L2 speakers; nS-Heb2, non-Semitic-L1 Hebrew-L2 speakers; S-Heb2; Semitic-L1 Hebrew-L2 speakers.

for Heb1 ($p < 0.001$) and S-Heb2 ($p = 0.005$) but not for nS-Heb2 ($p = 0.186$). In transposed, nS-Heb2 scored significantly higher than S-Heb2 ($p = 0.042$) and Heb1 ($p < 0.001$), with no difference between S-Heb2 and Heb1 ($p = 0.347$). No difference was shown in varying: S-Heb2 vs. nS-Heb2 ($p = 0.948$) and Heb1 ($p = 0.696$), and Heb1 vs. nS-Heb2 ($p = 0.535$).

These findings corroborate that when the consonantal roots are transposed, the Semitic language speakers, but not non-Semitic-L1, exhibit a significantly lower accuracy rate than varying consonantal roots irrespectively of grammatical information conveyed *via* the vocalic melody.

3.4. Rhyming HVM pairs

3.4.1. Binyanim relations

To examine whether awareness of the verbal binyanim relations affects the Hebrew-L2 speakers' language mode, we compared the HVM's three types, two of which express violation of syntactical relations (MIX, Pa'al-Pi'el, Pa'al-Pu'al) × language group (Heb1/Heb2). The analysis revealed a significant effect of language group [$F_{(1,327)} = 9.81$, $p = 0.002$] but not of binyanim relation types [$F_{(2,222)} = 0.49$, $p = 0.616$] and no interaction effect [$F_{(2,222)} = 0.20$, $p = 0.822$]. *Post hoc* analysis showed no difference between the three types for both language groups. However, Heb1 scored significantly lower than Heb2 in the MIX type ($p = 0.016$), which does not express binyanim relations, but not in the types that express binyanim relations Pa'al-Pi'el ($p = 0.157$) and Pa'al-Pu'al ($p = 0.110$).

Analysis discerning Heb2 by the participants' L1 [language group (Heb1/nS-Heb2/S-Heb2) × binyanim relation types (MIX, Pa'al-Pi'el, Pa'al-Pu'al)] revealed a significant effect of language group [$F_{(2,324)} = 19.00$, $p < 0.001$] but not of binyanim relation types [$F_{(2,221)} = 0.77$, $p = 0.463$] and no interaction effect [$F_{(4,221)} = 0.18$, $p = 0.950$]. *Post hoc* analysis showed that in all the HVM types, nS-Heb2 scored significantly higher than Heb1 and S-Heb2: MIX [Heb1 ($p = 0.001$), S-Heb2 ($p = 0.005$)], Pa'al-Pu'al [Heb1 ($p = 0.003$), S-Heb2 ($p = 0.001$)], and Pa'al-Pi'el [Heb1 ($p = 0.014$), S-Heb2 ($p = 0.007$)]. No differences were shown between Heb1 and S-Heb2 for all types: MIX ($p = 0.661$), Pa'al-Pi'el ($p = 0.289$), and Pa'al-Pu'al ($p = 0.152$) (Figure 2, asterisks in Data Table, left).

3.4.2. Semanitc relations

To further examine whether semantic relatedness in the HVM rhyming pairs affects language processing mode between Heb1 and Heb2, we compared language group (Heb1/Heb2) × semantic relatedness (Related/non-Related). The analysis revealed a significant effect of language group [$F_{(1,220)} = 7.45$, $p = 0.007$] but not of semantic relatedness [$F_{(1,220)} = 0.42$, $p = 0.517$] and no interaction effect [$F_{(1,220)} = 0.00$, $p = 0.995$]. *Post hoc* analysis showed that Related and non-Related pairs were similarly processed by Heb1 ($p = 0.644$) and Heb2 ($p = 0.649$). Differences between

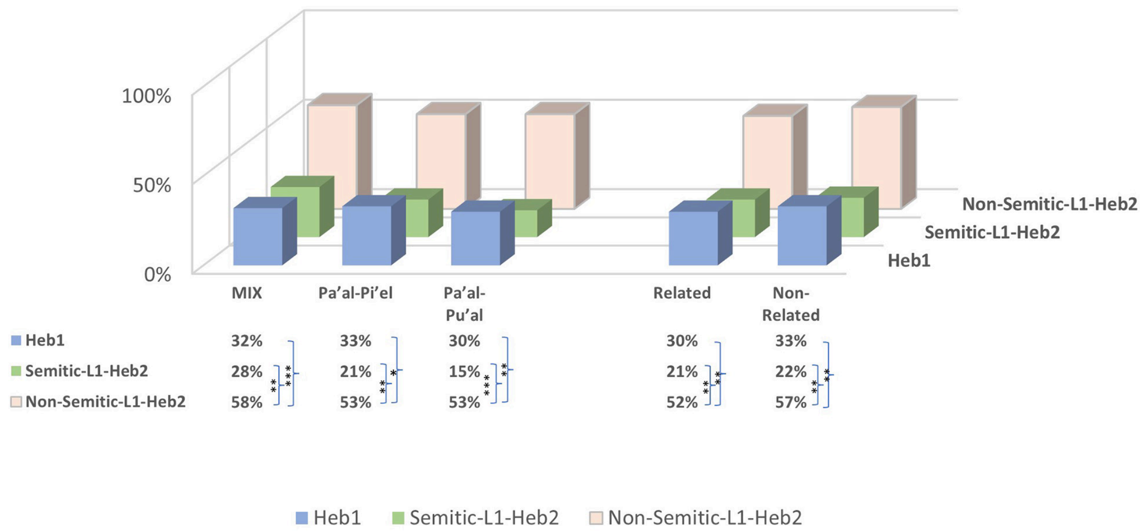


FIGURE 2
Binyanim relations and semantic relatedness in the HVM rhyming pairs. Asterisks in the data table indicate significant differences between language groups marked by the curly brackets according to conventional critical *P*-values: $*p < 0.05$, $**p < 0.01$, and $***p < 0.001$.

Heb1 and Heb2 approached significance in both Related ($p = 0.051$) and non-Related ($p = 0.061$) pairs.

Analysis discerning Heb2 by the participants' L1 [language group (Heb1/nS-Heb2/S-Heb2) × semantic relatedness (Related/non-Related)] revealed a significant effect of language group [$F_{(1,218)} = 13.83$, $p < 0.001$] but not of semantic relatedness [$F_{(1,218)} = 0.248$, $p = 0.619$] and no interaction effect [$F_{(2,218)} = 0.30$, $p = 0.970$]. *Post hoc* analysis showed that Related and non-Related pairs were similarly processed by all language groups: Heb1 ($p = 0.631$), S-Heb2 ($p = 0.959$), and nS-Heb2 ($p = 0.596$). By contrast, nS-Heb2 scored significantly higher than both Heb1 and S-Heb2 in both types: Related [Heb1 ($p = 0.002$), S-Heb2 ($p = 0.003$)] and non-Related [Heb1 ($p = 0.002$), S-Heb2 ($p = 0.002$)]. No differences were shown between Heb1 and S-Heb2: Related ($p = 0.384$) and non-Related ($p = 0.279$) (Figure 2, asterisks in Data Table, right).

These findings indicate that when the vocalic melody stands out (due to identical roots), especially but not exclusively, when binyanim relations are involved, and irrespectively of semantic relatedness, both Semitic language speakers exhibit a significantly low accuracy rate than non-Semitic-L1.

3.5. Non-stress-matched pairs

To examine the impact of stress on accuracy, we compared language group (Hbe1/Heb2) × non-stress-matched stimulus type [Stress/ Stress(−)]. The analysis showed significant effects of language group [$F_{(1,219)} = 13.44$, $p < 0.001$] and non-stress-matched stimulus type [$F_{(1,219)} = 4.64$, $p = 0.032$], with no interaction effect [$F_{(1,219)} = 0.17$, $p = 0.677$]. *Post hoc* analysis showed that accuracy rate was not different between Stress and Stress(−) for Heb1 ($p = 0.212$) and Heb2 ($p = 0.075$). However, Heb1 scored significantly higher than Heb2 in both Stress ($p = 0.006$) and Stress(−) ($p = 0.018$) (Table 5).

TABLE 5 Means and (SD) of non-stress-matched pairs.

	Heb1	Semitic-L1-Heb2	Non-Semitic-L1-Heb2	Heb2
Stress	0.77 (0.35) $_A$	0.69 (0.23)$_{ab}$	0.54 (39)$_b$	0.59 (0.35) $_B$
Stress (−)	0.85 (0.30) $_A$	0.85 (0.21)$_a$	0.64 (0.38)$_b$	0.70 (0.35) $_B$

Means in the same raw that do not share the sub-script are significantly different.
Lowercase indicates differences between Heb1, non-Semitic-L1-Heb2, and Semitic-L1-Heb2.
Uppercase indicates differences between Heb1 and Heb2.
Heb1, Hebrew native speakers; Heb2, Hebrew-L2 speakers.

The same analysis discerning Heb2 by the participants' L1 (language group (Heb1/nS-Heb2/S-Heb2) × non-stress-matched stimulus type [Stress/ Stress(−)] revealed significant differences in language group [$F_{(2,216)} = 9.89$, $p < 0.001$] and non-stress-matched stimulus type [$F_{(1,216)} = 4.73$, $p = 0.031$], with no interaction effect [$F_{(2,216)} = 0.179$, $p = 0.836$]. *Post hoc* analysis showed no significant differences between the two non-matched-stress types for all language groups: Heb1 ($p = 0.208$), nS-Heb2 ($p = 0.202$), and S-Heb2 ($p = 0.182$). However, Heb1 scored significantly higher than nS-Heb2 in both Stress ($p = 0.002$) and Stress(−) ($p = 0.002$), but not than S-Heb2: Stress ($p = 0.400$) and Stress(−) ($p = 0.971$). S-Heb2 scored significantly higher than nS-Heb2 in Stress(−) ($p = 0.034$) but not in Stress ($p = 0.182$) (Table 5).

These findings indicate that Semitic speakers are susceptible to phonological stress, whether the template is identical or different.

4. Discussion

This research explored cross-linguistic influence detached from communicative function by examining activation of language processing mechanisms in an auditory rhyming judgment task, which reflects on the transfer of (meta)linguistic information concerning Hebrew templatic words among Hebrew-L2 adult speakers with Semitic and non-Semitic-L1. The research included comparing Hebrew native speakers with Hebrew-L2 speakers in multi-layer analyses by the bilinguals' L1 due to different L1-L2 similarities and differences and two-resolution levels of transfer: the morphological processing and the sub-lexical morphemes and stress.

The findings show that in the control CVC pairs (not templatic words), accuracy was similarly high in rhyming pairs for all speakers, with no difference compared to non-rhyming pairs for Hebrew native speakers and non-Semitic-L1. Interestingly, Semitic-L1 scored lower in non-rhyming pairs compared to rhyming ones, suggesting an influence of rhyme perception of their mother tongue or insufficient phoneme discrimination in the L2. The results of the CVC pairs show that all language groups activated the linear processing mechanism, as required. The shift to CVCVC showed that all speakers scored lower in rhyming pairs, suggesting morphological processing. However, splitting the CVCVC pairs into stimulus types revealed that Semitic speakers (Hebrew native speakers and Semitic-L1) processed rhyming vs. non-rhyming pairs differently in all stimulus types but not non-Semitic-L1. This research corroborated the need for a multi-layer and high-resolution scrutiny, without which the results balanced out among Hebrew-L2 speakers and within stimulus types, giving an elusive impression of the appliance of the L2 mechanism.

Non-Semitic-L1, like Hebrew native speakers, processed rhyming vs. non-rhyming pairs differently in transposed pairs and in pairs where the vocalic melody is different (B(-) and HVM). However, unlike Hebrew native speakers, when the vocalic melody is identical [mVCCVC and B(+)], non-Semitic-L1 processed rhyming and non-rhyming pairs equally, resembling the processing of non-Hebrew speakers (Laure and Armon-Lotem, 2023b). Thus, in first resolution level, non-Semitic-L1 showed awareness of the morphological processing due to the vocalic melody template. By contrast, in second resolution level, the non-Semitic-L1's high accuracy rate in the HVM sub-type rhyming pairs was different from Semitic speakers, suggesting that lexical-syntactic linguistic information concerning the function of the vocalic melody was not strong enough to transfer. As expected, the phonological linguistic information concerning the root was not transferred. Comparing the varying vs. transposed phoneme rhyming pairs, no difference was shown for non-Semitic-L1 as opposed to Semitic speakers. Furthermore, accuracy was different in the non-stress-matched pairs (non-rhyming pairs) compared to the Hebrew native speakers, suggesting on face value that transfer of linguistic information about the contrastive stress did not occur. However, this language group comprises different L1s, and the results might have balanced, as contrastive stress is language-specific and changes also within language families, e.g., Spanish (contrastive) vs. French (non-contrastive). Together, these findings suggest that non-Semitic-L1 used their L1 processing mechanism with minor activation of L2 due to awareness of the morphological processing.

Semitic-L1 speakers, like Hebrew speakers, processed rhyming vs. non-rhyming pairs differently in transposed pairs and in pairs where the vocalic melody differs (B(-) and HVM). They also processed rhyming vs. non-rhyming pairs differently when the vocalic melody is identical (B(+), mVCCVC), but their results differed from Hebrew native speakers since accuracy was lower in the non-rhyming pairs than rhyming pairs and also differed from non-Semitic-L1, who showed no difference in rhyming vs. non-rhyming processing in these types. In rhyming pairs by stimulus types, Semitic-L1 showed a similar accuracy cascade $(\text{mVCCVC}=\text{B}(+)>\text{B}(-)>_{(\text{Heb1})}/=_{(\text{S}-\text{Heb2})}\text{TCR}>\text{HVM})$ to Hebrew native speakers. Thus, in first resolution level, Semitic-L1 showed awareness of the morphological processing. The results in the second resolution level confirm the awareness of the sub-lexical morphemes: Low accuracy rates were shown in transposed vs. varying consonantal roots and in the HVM, similar to those of Hebrew native speakers. In non-stress-matched pairs, Semitic-L1 also scored similarly to Hebrew native speakers despite the difference in stress between the languages, indicating activation levels of L2 and transfer of linguistic information. The low accuracy in non-rhyming pairs suggests insufficient phoneme discrimination or different rhyme perceptions compared to the Hebrew native speakers. Given that their phoneme discrimination was insufficient and yet their results resembled native speakers when the sub-lexical morphemes are accentuated corroborates their awareness of morphological processing and the sub-lexical morphemes, but was L1 or L2 the governing processing mechanism?

Although both could be equally applied, we find it L1 because of the low accuracy in non-rhyming pairs. The low accuracy could be due to rhyme perception. However, since the results differ from those of Hebrew speakers, it indicates that rhyme perception differs in these two languages, hence the governing mechanism was of L1. Another reason relates to phonological knowledge, which goes with phonetic knowledge expressed in phoneme perception and the ability to discriminate phonemes. Accuracy in non-rhyming pairs in CVC was relatively low, suggesting a non-native-like representation of Hebrew phonemes among the Semitic-L1. Transposed pairs accentuate the root for the phonological co-occurrence restrictions tapping into phonological computational knowledge based on phonemes and their distinguishing features. It is unlikely to activate phonological computational knowledge without native-like phoneme discrimination. Nevertheless, the Semitic-L1 showed accuracy similar to Hebrew native speakers in transposed vs. varying phonemes pairs. Therefore, the awareness of the phonological co-occurrence restrictions without an L2 phonemic representation similar (or close to similar) to Hebrew native speakers is likely to be filtered through L1, as the restrictions are common in Semitic languages.

Taken together, the findings indicate that Hebrew-L2 speakers processed templatic words activating their L1 governing processing mechanism, but not without activation of L2 mechanisms at different awareness levels despite the absence of the need to communicate or comprehend. This is not in accord with the unified competition model's (MacWhinney, 2005) premise that transfer of linguistic information and competition is for communicative function. Moreover, it emphasizes the impact of the non-linguistic parameters (Grosjean, 2001) of form, i.e., the modality, and function, i.e., participating in an experiment, on cross-linguistic influence. Stress is connected to the aural modality, and indeed awareness of L2 contrastive stress was transferred in Semitic-L1. The limitation of this study is its small sample size, which does not allow further investigation of the modality impact by breaking down the non-Semitic-L1 to each of the languages it contains. Furthermore, as we predicted, the awareness level played a role in the transfer. The vocalic melody template, which is linguistic knowledge taught or inducible based on usage, contributed to the awareness of morphological processing in non-Semitic-L1. This agrees with the unified competition model that associates linguistic transfer in adults with the necessity of linguistic awareness.

Interestingly, no activation of L2 was shown concerning the root phonological co-occurrence restrictions. One possible explanation is that, as suggested, this linguistic knowledge is subliminal and none of the non-linguistic parameters triggered its activation. Another explanation draws on the unified competition model, associating entrenchment with the strength of linguistic cues. Further research that includes balanced bilingual children may elucidate this subject. Also, investigation involving formal phonological and phonotactic theories may contribute to the understanding of the cross-linguistic influence concerning the root phonological restrictions in comparison with universal principles, including theories that are not syllable-dependent, like Net Auditory Distance (Dziubalska-Kołaczyk, 2014), that calculate phonemic distance based on their features and markedness.

To conclude, this study expands the scope of cross-linguistic influence research by investigating linguistic arenas when semantics is reduced to better understand how human language is processed. Activation of the L2 language mechanism without semantics projects on the brain plasticity and the neural circuits constructed due to bilingualism. The dynamic activation of all the bilingual languages, without context or semantic demands, enhances the benefit and contribution of bilingualism and cross-linguistic influence on the bilingual's linguistic toolbox. Of importance is the linguistic precision required for obtaining a better understanding of the complexity of the bilingual language mechanisms at work.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving human participants were reviewed and approved by the Ethics Committee of the Faculty of Humanities, Bar-Ilan University. The patients/participants provided their written informed consent to participate in this study.

Author contributions

YL: conceptualization, methodology, validation, data collection and curation, formal analysis, investigation, visualization, writing—original draft preparation, and reviewing and editing. SA-L: conceptualization, supervision, funding acquisition, and writing—reviewing and editing. Both authors contributed to the article and approved the submitted version.

Funding

YL was supported within ISF 863/14 and ISF 454/18.

Acknowledgments

We are in debt to Sean Adams for programming the online experiment and managing data extraction. We thank Yaniv Kanat-Maymon and Elad Shoham for statistical advice and Rachel Shoham for her comments.

Conflict of interest

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

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2023.1164510/full#supplementary-material

References

Aronoff, M. (1976). Word formation in generative grammar. Linguistic Inquiry Monographs Cambridge. *Mass* 1, 1–134.

Bat-El, O. (1994). Stem modification and cluster transfer in Modern Hebrew. *Nat. Lang. Linguistic Theory* 12, 571–596. doi: 10.1007/BF00992928

Bat-El, O. (2003). Semitic verb structure within a universal perspective. *Lang. Acquisit. Lang. Disord.* 28, 29–60. doi: 10.1075/lald.28.02bat

Benmamoun, E. (2003). The role of the imperfective template in Arabic morphology. *Lang. Acquisit. Lang. Disord.* 28, 99–114. doi: 10.1075/lald.28.05ben

Berent, I. (2017). "Algebraic phonology," in *The Routledge Handbook of Phonological Theory* (London: Routledge), pp. 569–588.

Bley-Vroman, R. (2018). Language as "something strange." *Bilingualism Lang. Cogn.* 21, 913–914. doi: 10.1017/S136672891800024X

Botwinik, I., Efrat, H., and Sharon, A. L. (2016). "Production of object relatives in bilingual acquisition: L1 Russian, L2 Hebrew," in *Bilingual Landscape of the Contemporary World*, eds Grucza, S., Olpińska-Szkiełko, M., Romanowski, P. (Frankfurt/Main - London - New York: Peter Lang Verlag).

Branum-Martin, L., Tao, S., and Garnaat, S. (2015). Bilingual phonological awareness: reexamining the evidence for relations within and across languages. *J. Educ. Psychol.* 107, 111. doi: 10.1037/a0037149

Degani, T., Prior, A., and Tokowicz, N. (2011). Bidirectional transfer: the effect of sharing a translation. *J. Cogn. Psychol.* 23, 18–28. doi: 10.1080/20445911.2011.445986

Deutsch, A., and Frost, R. (2003). Lexical organization and lexical access in a non-concatenated morphology. *Lang. Acquisit. Lang. Disord.* 28, 165–186. doi: 10.1075/lald.28.09deu

Dziubalska-Kołaczyk, K. (2014). Explaining phonotactics using NAD. *Lang. Sci.* 46, 6–17. doi: 10.1016/j.langsci.2014.06.003

Even-Shoshan (1974). *The New Dictionary*. Jerusalem: Kiryat-Sefer (in Hebrew).

Fox, B., and Routh, D. K. (1975). Analyzing spoken language into words, syllables, and phonemes: a developmental study. *J. Psycholinguist. Res.* 4, 331–342. doi: 10.1007/BF01067062

Gafni, C., Yablonski, M., and Ben-Shachar, M. (2019). Morphological sensitivity generalizes across modalities. *Mental Lexicon* 14, 37–67. doi: 10.1075/ml.18020.gaf

Greenberg, J. H. (1950). The patterning of root morphemes in Semitic. *Word* 6, 162–181. doi: 10.1080/00437956.1950.11659378

Grosjean, F. (1998). Studying bilinguals: methodological and conceptual issues. *Bilingual. Lang. Cogn.* 1, 131–149. doi: 10.1017/S136672899800025X

Grosjean, F. (2001). "The bilingual's language modes," in *One Mind, Two Languages: Bilingual Language Processing*, ed J. Nicol (Oxford: Blackwell), 1–22.

Heath, I. (2003). "Arabic derivational ablaut, processing," in *Language Processing and Acquisition in Languages of Semitic, Root-Based, Morphology*, ed J. Shimron (John Benjamins). p. 394.

Janssen, B., Meir, N., Baker, A., and Armon-Lotem, S. (2015). "On-line comprehension of Russian case cues in monolingual Russian and bilingual Russian-Dutch and Russian-Hebrew children," in *BUCLD 39: Proceedings of the 39th annual Boston University conference on language development* (Somerville, MA: Cascadilla Press), pp. 266–278.

Laure, Y., and Armon-Lotem, S. (2023a). *Linear and Nonlinear Processing of Hebrew templatic Words: The Role of Metalinguistic Awareness. [manuscript submitted for publication].*

Laure, Y., and Armon-Lotem, S. (2023b). *Language-universal vs. specific: non-Hebrew and Hebrew native speakers process auditory Hebrew templatic words. [manuscript submitted for publication].*

Lenel, J. C., and Cantor, J. H. (1981). Rhyme recognition and phonemic perception in young children. *J. Psycholinguist. Res.* 10, 57–67. doi: 10.1007/BF01067361

Lewkowicz, N. K. (1980). Phonemic awareness training: what to teach and how to teach it. *J. Educat. Psychol.* 72, 686. doi: 10.1037/0022-0663.72.5.686

MacWhinney, B. (2005). A unified model of language acquisition. *Handbook Bilingualism: Psycholinguist. Approach.* 4967, 50–70. doi: 10.1184/r1/6613271.v1

MacWhinney, B. (2013). "The logic of the unified model," in *The ROUTLEDGE Handbook of Second Language Acquisition* (London: Routledge), pp. 229–245.

Marslen-Wilson, W., Tyler, L. K., Waksler, R., and Older, L. (1994). Morphology and meaning in the English mental lexicon. *Psychologic. Rev.* 101, 3. doi: 10.1037/0033-295X.101.1.3

McCarthy, J. (1981). A prosodic theory of nonconcatenative morphology. *Linguistic Inq.* 12,. 373–418.

Meir, N., Walters, J., and Armon-Lotem, S. (2017). Bi-directional cross-linguistic influence in bilingual Russian-Hebrew children. *Linguistic Approach. Bilingualism* 7, 514–553. doi: 10.1075/lab.15007.mei

Norman, T., Degani, T., and Peleg, O. (2016). Transfer of L1 visual word recognition strategies during early stages of L2 learning: evidence from Hebrew learners whose first language is either Semitic or Indo-European. Second *Language Res.* 32, 109–122. doi: 10.1177/0267658315608913

Norman, T., Degani, T., and Peleg, O. (2017). Morphological processing during visual word recognition in Hebrew as a first and a second language. *Readi. Writ.* 30, 69–85. doi: 10.1007/s11145-016-9663-7

Oganyan, M., Wright, R., and Herschensohn, J. (2019). The role of the root in auditory word recognition of Hebrew. *Cortex* 116, 286–293. doi: 10.1016/j.cortex.2018.06.010

Segal, O., and Kishon-Rabin, L. (2019). Influence of the native language on sensitivity to lexical stress: evidence from native Arabic and Hebrew speakers. *Stud. Second Lang. Acquisit.* 41, 151–178. doi: 10.1017/S027226311 7000390

Upasana, N., Stav, E., Brianna, L., Yamasaki, B. N., Vedran, D., et al. (2022). "Learning and consolidation of morphologically derived words in a novel language: evidence from Hebrew speakers," in *Society for the Neurobiology of Language, 2022 annual meeting, Philadelphia*.

Ussishkin, A. (2000). "Root-and-pattern morphology without roots or patterns," in *North East Linguistics Society* (Vol. *30, No.* 2, p. 18).

Ussishkin, A. (2003). Templatic effects as fixed prosody: the verbal system in Semitic. *Amsterdam Stud. Theory Hist. Linguistic Sci. Series* 4, 511–530. doi: 10.1075/cilt.241.23uss

Velan, H., and Frost, R. (2007). Cambridge University vs. Hebrew University: the impact of letter transposition on reading English and Hebrew. *Psychonomic Bull. Rev.* 14, 913–918. doi: 10.3758/BF03194121

Velan, H., and Frost, R. (2011). Words with and without internal structure: what determines the nature of orthographic and morphological processing?. *Cognition* 118, 141–156. doi: 10.1016/j.cognition.2010.11.013

Yablonski, M., and Ben-Shachar, M. (2016). The morpheme interference effect in hebrew: a generalization across the verbal and nominal domains. *Mental Lexicon* 11, 277–307. doi: 10.1075/ml.11. 2.05yab

Zion, D. B., Nevat, M., Prior, A., and Bitan, T. (2019). Prior knowledge predicts early consolidation in second language learning. *Front. Psychol.* 10, 2312. doi: 10.3389/fpsyg.2019.02312

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 17 August 2023
DOI 10.3389/fpsyg.2023.1224830

Check for updates

OPEN ACCESS

EDITED BY
Katarzyna Jankowiak,
Adam Mickiewicz University, Poland

REVIEWED BY
Malgorzata Kul,
Adam Mickiewicz University, Poland
Honghao Ren,
Xi'an Jiaotong University, China

*CORRESPONDENCE
Xiaodong Fei
✉ feixiaodonghi@bfsu.edu.cn

RECEIVED 18 May 2023
ACCEPTED 31 July 2023
PUBLISHED 17 August 2023

CITATION
Geng Y, Song Q and Fei X (2023) Factors in cognitive processing of Japanese loanwords by advanced Chinese Japanese-as-a-foreign-language learners.
Front. Psychol. 14:1224830.
doi: 10.3389/fpsyg.2023.1224830

Factors in cognitive processing of Japanese loanwords by advanced Chinese Japanese-as-a-foreign-language learners

Yaoyao Geng[1], Qichao Song[1] and Xiaodong Fei[2]*

[1]Graduate School of Humanities and Social Sciences, Hiroshima University, Hiroshima, Japan, [2]Beijing Center for Japanese Studies, Beijing Foreign Studies University, Beijing, China

Introduction: Previous studies have highlighted the challenges faced by Chinese Japanese-as-a-foreign-language (JFL) learners (whose L2 is English) in acquiring L3 Japanese loanwords. These challenges arise from the linguistic characteristics of loanwords and the limited emphasis on teaching and learning them. However, there is a lack of research on the specific factors that influence the processing of Japanese loanwords among Chinese JFL learners. Significant motivation exists, therefore, to investigate these influencing factors as they provide valuable insight into the integration of phonographic and ideographic language systems, ultimately facilitating future lexical acquisition.

Methods: In this study, an experiment was conducted on 31 Chinese JFL learners to investigate the effects of loanword familiarity, English vocabulary proficiency, English-Japanese phonological similarity, and context on the processing of Japanese loanwords.

Results: Data analysis, using a (generalized) linear mixed-effect model, provided the following insights: (1) the processing of Japanese loanwords is influenced by English-Japanese phonological similarity, loanword familiarity, context, and learner English proficiency. Among these four factors, familiarity has the most significant impact on Japanese loanword processing; (2) the effects of context and phonological similarity on the processing of Japanese loanwords are not consistently positive. As learners improve their proficiency in L3 Japanese, they tend to decrease their reliance on English knowledge and instead access loanword representations directly to conceptual representations.

Discussion: Based on the findings of this study, a processing model for Japanese loanwords among advanced Chinese JFL learners is proposed. The model emphasizes the critical importance of the characteristics of loanwords, including phonological similarity and familiarity. It is necessary to determine the specific circumstances in which context considerably enhances learner processing ability.

KEYWORDS

Japanese loanwords processing, English-Japanese phonological similarity, familiarity, context, English vocabulary proficiency, Chinese Japanese-as-a-foreign-language learners

1. Introduction

Lexical processing in second languages (L2s), mainly in phonographic languages, has received considerable attention lately. Understanding how L2 learners store and retrieve words provides valuable insight into lexical acquisition and processing (Lee et al., 2018; Jankowiak, 2021). In recent years, there has been a surge in studies focusing on the processing of Japanese *Kanji* words by Chinese Japanese-as-a-foreign-language (JFL) learners. This has provided fresh empirical evidence in the field of ideographic writing systems (e.g., Mori, 2014; Fei and Li 2017; Fei et al., 2022; Song et al., 2023). Japanese writing systems can be divided into three types: *Kanji, Hiragana, and Katakana*. These systems encompass both phonographic and ideographic elements. *Kanji* originated from Chinese characters and is used by both Chinese and Japanese. *Hiragana* is used for native Japanese words and grammatical elements, whereas *Katakana* is primarily used for loanwords, which are commonly referred to as 外来語 "*Gairaigo*" or カタカナ語 "*Katakanago*" (Kess and Miyamoto, 2000). With the rapid advancement of informatization and globalization, the number of Japanese loanwords is increasing. According to Sube (2013), approximately 80 percent of the loanwords listed in the Iwanami Kokugo Jiten (Iwanami Japanese Dictionary, 3rd Edition) are derived from English. In China, there are millions of Japanese language learners, they are number second only to the number of English learners. Investigating the factors that influence the processing of loanwords by Chinese JFL learners can provide valuable insight into the integration of phonographic and ideographic language systems for future lexical acquisition research.

Lexical proficiency encompasses two dimensions: vocabulary breadth and vocabulary depth (Wesche and Paribakht, 1996; Li and Kirby, 2015), which relate to the relationship between quantity and quality. As learners progress to an advanced stage in their language-learning journey, when their vocabulary breadth reaches a certain level, how quickly they process and retrieve existing vocabulary from their mental lexicon becomes increasingly important. Therefore, it is necessary to explore the factors that affect the lexical processing which facilitates this process. Nevertheless, as loanwords are an important component of Japanese vocabulary, little research has been conducted on the processing of loanwords by Chinese JFL learners compared with research on Japanese *Kanji* word processing (Tamaoka, 1997; Yamato et al., 2010; Yamato and Tamaoka, 2011, 2013; Jha et al., 2018; Tamaoka, 2018).

To fill this research gap, the present study investigates factors that influence the processing of loanwords by advanced Chinese JFL learners. It examines the influence of English-Japanese phonological similarity, familiarity, context, and English vocabulary proficiency on the processing of Japanese loanwords.

2. Literature review

2.1. Hypotheses on an L2 lexical processing model

Numerous studies have provided substantial evidence supporting the phenomenon of "non-selective processing" in bilinguals. Results suggest that, when processing one language, bilinguals unintentionally activate both the conceptual and lexical representations of another (Hermans et al., 1998; De Groot et al., 2000; Van Hell and Dijkstra, 2002; Singh et al., 2014; Dijkstra and Walter, 2018). The effects of orthographic, phonological, and semantic similarities have been widely documented in both first language (native language, L1) and L2 contexts (Antón and Duñabeitia, 2020). Nonetheless, whether these effects also exist when processing L2s and L3s simultaneously is unclear.

Previous studies (e.g., Dewaele, 1998; De Angelis and Selinker, 2001; Jasone, 2001) have indicated that the language knowledge acquired in one language can affect the processing of another language, particularly in the case of L2 learners. Chinese JFL learners possess knowledge not only of their own Chinese (L1), but also of English (L2), which they typically study in school to prepare for university entrance exams. As a result, Chinese JFL learners are likely influenced by their knowledge of English when acquiring Japanese (L3). Additionally, the Japanese language incorporates a considerable number of loanwords derived from English, which are written in *Katakana* using English pronunciation. This means that there are numerous English-derived loanwords that exhibit high phonological similarity to English, enabling Chinese JFL learners to draw on their English knowledge when processing Japanese loanwords (Hoshino and Kroll, 2008; Yamato and Tamaoka, 2013). Therefore, investigating the impact of English (L2) on the processing of Japanese (L3) loanwords by Chinese JFL learners offers a valuable approach to understanding the interaction between languages and their influence on language processing.

The Revised Hierarchical Model (Figure 1) proposed by Kroll and Steward (1994) has been widely employed in bilingual lexical processing research (e.g., Silverberg and Samuel, 2004; Ferré et al., 2006; Kroll et al., 2010). According to this model, links between lexical and conceptual representations in a target language are believed to develop as learner proficiency in the target language improves. As L2 proficiency increases, reliance on the L1 for conceptual links gradually diminishes, particularly when dealing with two languages from different language families. Numerous empirical studies have been conducted on *Kanji* word processing among Chinese JFL learners (e.g., Matsumi et al., 2012; Fei et al., 2022; Song et al., 2023). During the initial stages of Japanese lexical acquisition, links between L3 Japanese lexical and conceptual representations tend to be weak. Therefore, the processing of *Kanji* word is expected to rely on

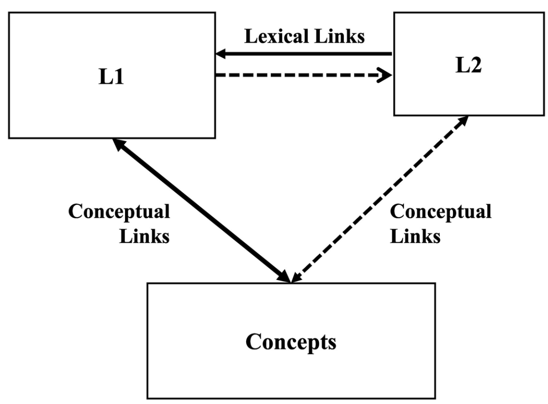


FIGURE 1
Revised hierarchical model (Kroll and Steward, 1994).

translation equivalents in the learner's L1 (i.e., Chinese), which serves as a bridge to accessing corresponding conceptual representations. In contrast, an advanced learner's links between Japanese lexical and conceptual representations tend to be strong, facilitating his or her direct access to conceptual representations without relying heavily on L1 knowledge. Japanese (L3) loanwords are related semantically to the Chinese (L1), and phonetically to English, the acquired L2. Therefore, unlike Chinese-Japanese bilingual processing in *Kanji* words, processing models of loanwords may involve a complex trilingual interaction in terms of phonetics and semantics. Nonetheless, as mentioned above, there is limited research and theoretical discussion on the construction of such processing models and their theoretical implications.

2.2. Factors affecting the processing of Japanese loanwords

Previous research has focused on the processing of Japanese loanwords in the context of bilingualism, specifically examining how it is influenced by English-Japanese phonological similarity and learners' English proficiency (Yamato and Tamaoka, 2013; Tamaoka, 2018). Moreover, studies have highlighted the significance of other influencing factors, including familiarity and the presence or absence of context, in shaping the processing of Japanese loanwords (Geng, 2022).

As mentioned earlier, the processing of loanwords is influenced by English-Japanese phonological similarity and the English proficiency of learners because many Japanese loanwords are derived from English. In contrast to English syllables, which display considerable structural variation and can be quite complex, there are only four fundamental syllable structures in Japanese: (C) V (C = consonant; V = vowel), (C) VV, (C) VN (N = nasal), and (C) VQ (Q = first part of a geminate obstruent) (Tajima et al., 2002). Therefore, loanwords undergo phonological assimilation to conform to the primarily CV-based structure of Japanese. For example, "テキスト," borrowed from the word "text," is pronounced as /tekisuto/, and "ドリーム," borrowed from the word "dream" is pronounced as /dori:mu/. Tamaoka (2018) conducted a priming experiment and found that loanword processing was facilitated when there was a high degree of phonological similarity with English. Tamaoka (1997) also found that, compared with those of English JFL learners, the accuracy rates of Chinese JFL learners were significantly lower. This suggests that English vocabulary proficiency plays an important role in accurately processing Japanese loanwords. Yamato and Tamaoka (2013) further investigated the influence of English knowledge on the processing of loanwords. They revealed that English proficiency, as an L2, significantly influenced the processing of loanwords in Japanese as an L3. However, the Japanese proficiency levels of the participants of Yamato and Tamaoka (2013) range from intermediate to advanced. For Chinese learners, the level of English-Japanese phonological similarity is directly related to the number of English and Japanese words stored in their mental lexicon. Therefore, when examining the impact of English on the processing of Japanese loanwords, it is crucial to consider both phonological similarity and L2 and L3 proficiency. However, the research described above does not provide an in-depth exploration of the relationship between the influence of phonological similarity and language proficiency.

Research has consistently demonstrated the significant role of context in L2 lexical processing (Stanovich and Richard, 1983; Balota et al., 1985; Sereno et al., 2003; Goldwater et al., 2009; Perea et al., 2013). A recent study conducted by Song and Fei (2022) highlighted the substantial impact of context on the processing of Japanese vocabulary by Chinese JFL learners. Nevertheless, the role of context in facilitating loanword comprehension has not been confirmed. Jha et al. (2018) discussed the processing of novel words written in *Katakana* by advanced Chinese JFL learners and found no such significant role involving context. However, it is important to note that the stimuli used in Jha et al. (2018) were non-words written in *Katakana*, which may have caused participants to focus more on processing the meaning of the target non-words rather than comprehending the sentence as a whole. Thus, further investigation is necessary to gain a clearer understanding of how context influences the processing of loanwords by Chinese JFL learners. Familiarity is also an important factor in influencing a learner's processing of Japanese loanwords. Yamato and Tamaoka (2011) found that for Chinese JFL learners with low proficiency in Japanese, low-familiarity loanwords significantly influenced processing speed. Yamato et al. (2010) suggested that, for Chinese JFL learners, the processing efficiency of loanwords was influenced to a large extent by learning duration. Specifically, as the learning duration increases, their familiarity with Japanese loanwords also improves. Consequently, the activation threshold for words decreased, allowing for faster activation of representations. Additionally, a previous study has revealed that, in advanced Chinese EFL (English-as-a-foreign-language) learners, processing patterns are still influenced by English word familiarity (Li et al., 2011). This suggests that, if bilingual individuals are highly familiar with L2 vocabulary, they can directly access concepts from the L2 (Chen and Leung, 1989). Consequently, further research is needed to explore the influence of familiarity on the processing of loanwords among Chinese JFL learners as their Japanese proficiency improves.

In summary, there is currently a lack of research on the various factors that influence the processing of Japanese loanwords. Existing studies have focused on bilingual proficiency, phonological similarity, and context, without fully examining their interactions and specifically emphasizing advanced Chinese JFL learners. Therefore, research on the influence of these factors is urgently required to elucidate this phenomenon.

2.3. Objectives and hypotheses of this study

This study investigates the influence of English-Japanese phonological similarity, familiarity, context, and English vocabulary proficiency on the processing of Japanese loanwords in advanced Chinese learners. Motivation for the study rests on the lack of comprehensive research that systematically investigates the various influencing factors, and the learner's Japanese proficiency not being restricted to one level only. Whether such influencing factors undergo changes during processing with increasing learner language proficiency is investigated. This study focuses on the following questions:

RQ1: How do English-Japanese phonological similarity, familiarity, context, and English vocabulary proficiency influence the accuracy rates and reaction times of loanword processing among advanced Chinese JFL learners?

RQ2: What lexical processing models do advanced Chinese JFL learners utilize when processing Japanese loanwords?

Based on the review of existing studies given above, the hypotheses of this study are as follows:

H1: Previous research has demonstrated that phonological similarity enables Chinese JFL learners to rely on the pronunciation of corresponding English words in their mental lexicon (Yamato and Tamaoka, 2013; Tamaoka, 2018), facilitating the processing of Japanese loanwords. The current study expects English-Japanese phonological similarity to promote the processing of Japanese loanwords.

H2: Previous research (Li et al., 2011) shows that even advanced learner processing is influenced by familiarity. Therefore, it is speculated that, in the processing of Japanese loanwords by advanced Chinese JFL learners, both the accuracy rates and reaction times will be positively influenced by familiarity with Japanese loanwords.

H3: Concerning context, although Jha et al. (2018) found that the presence or absence of context did not influence the inference of unknown loanword meanings, considering their experimental materials were non-words and considering the numerous findings from previous research on language processing that demonstrate the facilitating role of context in lexical processing, it is speculated that context can play a facilitating role (e.g., Sereno et al., 2003; Goldwater et al., 2009). Specifically, we anticipate that both the accuracy rates and reaction times will improve.

H4: Previous research has confirmed that a high English vocabulary proficiency is associated with highly efficient processing of Japanese loanwords (Tamaoka, 1997; Yamato and Tamaoka, 2013). Therefore, it is speculated that English vocabulary proficiency may facilitate the processing of Japanese loanwords.

3. Materials and methods

3.1. Participants

An experiment was conducted on 31 advanced Chinese JFL learners, comprising 19 females and 12 males, with ages ranging from 22 to 26 years old. The participants had a mean Japanese study time of 6.02 ($SD = 1.42$) years. And all were enrolled in the same graduate school in China, majoring in Japanese language and literature. They began studying Japanese in their first year of college and passed the Japanese-Language Proficiency Test (JLPT) at the N1 level (the highest level, which, according to JLPT's official instructions, means they obtained the ability to understand Japanese in various circumstances). All participants had normal vision (with corrected vision). The participants, therefore, belonged to a homogeneous group of learners. We provided them with the Language History Questionnaire (Li et al., 2020) to assess the participants' proficiency and usage time in Chinese, Japanese, and English. Analysis of the questionnaire responses revealed that all participants were unbalanced trilinguals, with their highest proficiency in Chinese, followed by Japanese and English [*Fs* $(2, 60) = 88.23–193.90$, *ps* < 0.001, see Table 1].

TABLE 1 Participants' self-reported language proficiency and comparisons between Chinese, Japanese, and English.

	C	J	E	Comparison
L	6.71 (0.53)	4.90 (0.70)	3.19 (0.95)	C > J > E***
S	6.42 (0.85)	4.65 (0.84)	2.77 (1.20)	C > J > E***
R	6.58 (0.96)	5.58 (0.99)	3.81 (1.38)	C > J > E***
W	6.16 (1.13)	4.81 (1.01)	2.90 (1.22)	C > J > E***
Time of Usage (h/day)	11.78 (3.96)	5.98 (3.45)	1.02 (0.94)	C > J > E***

****p* < 0.001; L, listening; S, speaking; R, reading; W, writing; C, Chinese; J, Japanese; E, English. *Tukey's HSD tests* were used for multiple comparisons.

3.2. Design

In this study, the impact of four independent variables on the accuracy rate and reaction time of Chinese JFL learners was examined. These variables were English-Japanese phonological similarity, familiarity with loanwords, context and English vocabulary proficiency.

3.3. Materials

Forty-four word items and 22 sentences for contextual condition (see supplementary materials) were created. The selection of loanwords was based on the list of Basic Loanwords List by Mochizuki (2012) and Japanese textbooks (Peng and Moriya, 2007) used in Chinese universities. To ensure an appropriate level of difficulty in the loanword materials, we utilized "Reading Tutor,[1]" a widely recognized website for Japanese education research that assesses the difficulty of Japanese content. The difficulty of the loanwords was adjusted based on the analysis results from "Reading Tutor," resulting in a word list consisting of 90 English-derived loanwords. The control procedures for the various indicators of the experimental materials are described as follows.

[Phonological Similarity] The 90 loanwords were recorded in both standard English and Japanese by a native English speaker from England and a native Japanese speaker from Japan. Due to the relatively high familiarity of Japanese loanwords among Chinese JFL learners, especially among advanced learners, there may have been a bias in their perception of the phonological similarity between English and Japanese loanwords. Therefore, to specifically examine the phonological similarity for Chinese learners, we recruited 20 Chinese university students who had prior English learning experience and

1 https://chuta.cegloc.tsukuba.ac.jp/

achieved level 4 in the College English Test, a well-known English proficiency test held in China annually to test the English proficiency of Chinese university students. These participants had no prior experience in learning Japanese. They were assigned a phonological similarity judgment task using a seven-point rating questionnaire. The rating scale ranged from 1 (not similar at all) to 7 (very similar).

[Familiarity] For the familiarity evaluation of the selected materials, we recruited 76 Chinese JFL learners who had a background in Japanese learning similar to that of the participants in the experiment. They were instructed to rate the materials on a seven-point scale, ranging from 1 (not familiar at all) to 7 (very familiar). The reason we used a seven-point rating questionnaire rather than a five-point one was because phonological similarity and familiarity were used as continuous variables when conducting data analysis. Therefore it is ideal to have sufficient statistical dispersion to ensure a more significant linear change between familiarity and phonological similarity judgment values.

[Context] To ensure differentiation between high and low phonological similarity and familiarity in the experimental materials, we carefully selected 44 loanwords from the initial pool of 90. The selection was based on the evaluation results of phonological similarity and familiarity obtained from previous assessments. These loanwords were then equally divided into two groups, an isolated condition group and a contextual condition group, with 22 words assigned to each condition. During the selection process, we took into account various factors, such as word frequency (based on the Balanced Corpus of Contemporary Written Japanese, BCCWJ), mora number, phonological similarity, and familiarity, and ensured that there were no significant differences in the above indicators between the two sets of materials [$ts\ (42) = 0.06\text{–}1.33$, $ps > 0.192$, see Table 2].

Contextual sentences containing the target loanwords were carefully selected from the BCCWJ and online sources. The selected contexts were then evaluated in terms of their degree of matching with the loanwords. To ensure the highest level of relevance between the contextual materials and loanwords, we invited five advanced Chinese JFL learners to reconfirm the suitability of the chosen contexts. Additionally, we utilized the "Reading Tutor" tool to evaluate the level of difficulty of the materials used in our experiment, ensuring that the chosen contexts did not hinder participants' smooth reading. The results from the "Reading Tutor" indicated that all sentences were considered "very easy." Therefore, the level of difficulty of the contexts did not impact participant processing.

[English Vocabulary Proficiency] To assess the participants' English vocabulary proficiency, we utilized the Bilingual Version of Vocabulary Size Test (VST) developed by Nation and Beglar (2007). This modified version of the Vocabulary Level Test was created based on Nation (1983). The VST focuses on evaluating learners' receptive vocabulary and consists of 14 sections, each representing a different vocabulary level ranging from 1,000 to 14,000 words. Each section comprises 10 multiple-choice questions, featuring an English word, a contextless sentence, and four Chinese semantic alternatives. Participants are required to choose the correct answer from the four options, and they receive one point for each correct response. This test has been widely used to assess the English writing vocabulary of L2 learners due to its reliability and validity (Nation and Beglar, 2007).

[Fillers] In addition to the 44 selected words for the lexical judgment task, 28 non-words (14 for the isolated condition group and 14 for the contextual condition group) as fillers were selected. These non-words consisted of two types: those that were similar to the original words, such as "バイオリーン" (correct loanword: "バイオリン," violin), and those that were dissimilar, such as "サドバハヤ." The sentences used for fillers in the contextual condition group were collected from the BCCWJ and Japanese textbooks mentioned above. The "Reading Tutor" showed that all sentences were "very easy."

3.4. Apparatus

A personal computer (SOTEC N15 WMT02) was used for the loanwords' presentation. The experimental program was created using SuperLab Pro 4.0 (Cedrus Corporation).

3.5. Procedure

Because of the potential limitations associated with using reaction time as a measure in psychological experiments (see Crocetta and Andrade, 2015), we carefully selected the experimental apparatus and created a controlled environment for implementation. Participants were tested individually in a sound-attenuated room. Figure 2 shows the experimental procedure for one trial under both isolated and contextual conditions. Under the isolated condition, each trial began with the display of a fixation point on the screen for 500 ms, indicating the upcoming appearance of a loanword. After a 200 ms blank interval, a loanword was presented on the screen. The maximum presentation time for each word was 5,000 ms. Once a participant responded, or if 5,000 ms elapsed without a response, the next trial commenced after 2000 ms. The stimuli were presented in randomized order. In the contextual condition group, a sentence was presented before a loanword was presented. The sentence contained a blank space, and participants were instructed to read and understand the sentence and judge whether the upcoming loanword existed in Japanese. Except for the inclusion of sentences before participants make judgments, the procedure for presenting stimuli was the same in both the isolated and contextual condition groups.

TABLE 2 Summary of characteristics of the test items.

	Mora count	Logged-transformed frequency	Phonological similarity	Familiarity	Examples
Isolated condition	3.45 (0.80)	3.16 (0.51)	4.25 (1.67)	5.81 (0.97)	プラス (plus)
Contextual condition	3.86 (1.21)	3.06 (0.52)	4.28 (1.70)	5.89 (1.30)	台風のせいでテレビのアンテナが折れてしまいました。(Due to the typhoon, the TV antenna got broken)

Results are expressed as mean (*SD*).

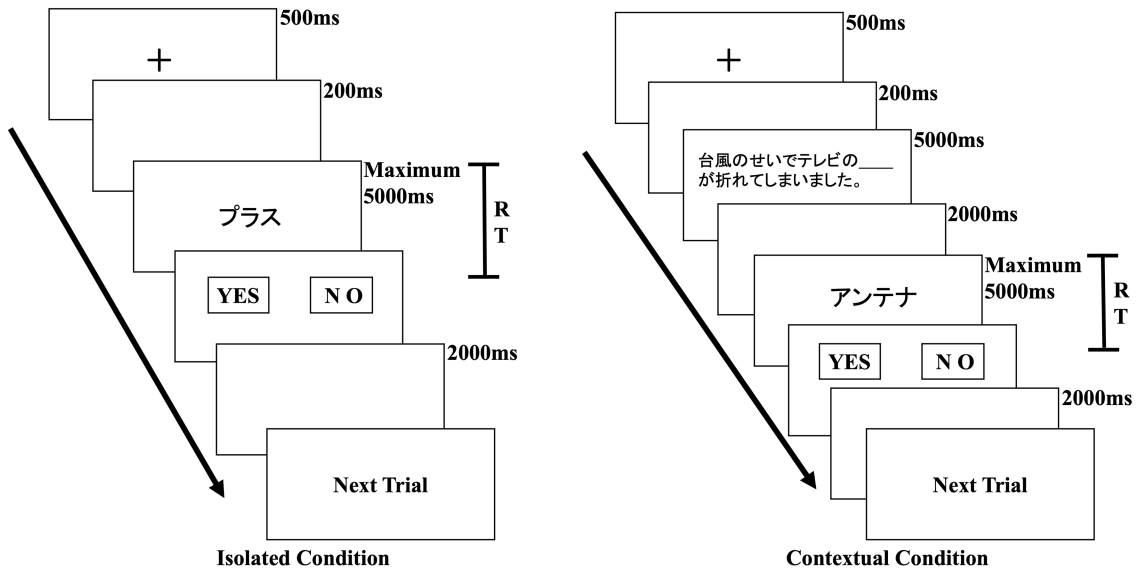


FIGURE 2
Flow of one experimental trial under isolated and contextual conditions.

TABLE 3 Results of GLME model analysis of accuracy rates.

Variables	Estimate	SE	*z*	*pr* (>|*z*|)
Intercept	9.22	2.48	3.72	<0.001
Familiarity	3.19	1.51	2.11*	0.035
Phonological Similarity	3.06	1.38	2.22*	0.027
ContextY	−3.51	2.93	−1.20	0.231
Phonological Similarity: ContextY	−9.24	3.09	−2.99**	0.003
Familiarity: Phonological Similarity: ContextN	1.26	1.11	1.14	0.256
Familiarity: Phonological Similarity: ContextY	−7.03	1.74	−4.04***	<0.001

*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.
Participants = 31. Items = 44. Total observation = 1181.
SE, standard error. *df*, degree of freedom.
The optimal model is *glmer* [acc ~ Familiarity + Phonological Similarity + Context + Phonological Similarity: Context + Familiarity: Phonological similarity: Context + (1|item) + (1|participant), data = data, glmerControl optimizer = "bobyqa," optCtrl = list (maxfun = 100000)].

Participants were given instructions to determine, as quickly and accurately as possible, whether each stimulus constituted a genuine word. Before the formal experiment began, we conducted sufficient practice and testing sessions to ensure that the participants fully understood the experimental task.

4. Results

4.1. Data manipulation

We excluded 23 out of the trials whose reaction times were longer than 3,500 ms and ±2.5 *SD*s above and below the mean. The percentage of exclusion was 1.70%. To deal with skewed data, reaction times were log-transformed. The phonological similarity and English vocabulary proficiency data were standardized. Data analyses were conducted using the software R (version 4.2.1, R Core Team, 2022). We used linear-mixed-effect model with the lme4 (Bates et al., 2015) and lmerTest (Kuznetsova et al., 2017) packages. The model with the lowest Akaike information criterion (AIC) was selected as the optimal model for model fitting. The software jamovi (version 2.3, jamovi project, 2022) was used to examine interactions. The Wald *z*-distribution was used to compute *p*-values for the accuracy rates data. The Wald *t*-distribution approximation was used to compute *p*-values for the reaction times data.

4.2. Results of the accuracy rates data

Using the AIC, familiarity, context, phonological similarity, first-order interaction of similarity and context, and second-order interaction of similarity, the context and familiarity were selected as fixed effects, and participants and items were selected as random effects in the model. The results for the accuracy rates are shown in Table 3. The main effect of familiarity was significant (χ^2 (1) = 4.45, $p = 0.035$), indicating that the accuracy rate increased with increasing familiarity. The main effect of phonological similarity was significant (χ^2 (1) = 4.91, $p = 0.027$), indicating that the accuracy rate increased with increasing phonological similarity. The main effect of context was not significant (χ^2 (1) = 1.43, $p = 0.231$).

 10.3389/fpsyg.2023.1224830

Given the significant first-order interaction between phonological similarity and context (χ^2 (1) = 8.95, $p=0.003$), simple main effects were analyzed (Table 4). The results indicate that, under the isolated condition, loanwords with high phonological similarity had significantly higher accuracy rates than those with low phonological similarity ($z=2.67$, $p=0.008$). In contrast, under the contextual condition, loanwords with high phonological similarity had lower accuracy than those with low phonological similarity ($z=-2.72$, $p=0.007$). Additionally, when phonological similarity was low, accuracy tended to be higher with context than without context ($z=-1.76$, $p=0.078$). However, when phonological similarity was high, accuracy was significantly lower with context than without context ($z=2.51$, $p=0.012$).

Furthermore, there was a significant second-order interaction between context, phonological similarity, and familiarity (χ^2 (2) = 17.09, $p<0.001$). The results for the simple main effects are shown in Table 5. When both familiarity and phonological similarity are high, the context has an inhibitory effect. In contrast, when both familiarity and phonological similarity are low, despite the relatively low accuracy rates in the presence of context, there is not a significant effect from the context.

TABLE 4 Results of simple main effects between context and phonological similarity.

Moderator Levels	Contrast	Estimate	SE	*z*	*pr* (>|*z*|)
N	High – Low	7.69	2.88	2.67**	0.008
Y	High – Low	−1.56	0.57	−2.72**	0.007
Low	N – Y	−5.73	3.25	−1.76†	0.078
High	N – Y	12.76	5.07	2.51*	0.012

†$p<0.10$, *$p<0.05$, **$p<0.01$.
N, isolated condition. Y, contextual condition.
SE, standard error; *df*, degree of freedom.
High, Mean + *1SD*. Low, Mean-*1SD*.

4.3. Results of the reaction times data

Only correct responses to Yes trials were included in the analysis. Using the AIC, familiarity, English vocabulary proficiency, context, and first-order interaction of English vocabulary proficiency and context were selected as fixed effects, and participants and items were selected as random effects in the model. The results for the reaction times are shown in Table 6. The main effect of familiarity was significant (F (1, 45.44) = 38.47, $p<0.001$), indicating that participants responded faster with increasing levels of familiarity. The main effect of English vocabulary proficiency was non-significant (F (1, 29.01) = 0.05, $p=0.816$). Similarly, the main effect of context did not reach statistical significance (F (1, 41.07) = 2.26, $p=0.141$). However, a significant interaction was observed between English vocabulary proficiency and context (F (1, 1108.40) = 4.01, $p=0.046$).

Table 7 presents results from the simple main effect tests. English vocabulary proficiency had no significant effect whether or not there was context [t (32.17) = 0.22, $p=0.831$; t (32.05) = 0.67, $p=0.506$]. For learners with high English vocabulary proficiency, there was no significant difference between the contextual and isolated conditions [t (55.41) = 0.65, $p=0.519$]. However, for learners with low English vocabulary proficiency, the reaction time under the contextual condition was significantly shorter than that under the isolated condition [t (55.20) = 2.14, $p=0.037$].

TABLE 5 Results of simple main effects between context, phonological similarity, and familiarity.

Moderator Levels						
Phonological similarity	**Familiarity**	**Contrast**	**Estimate**	**SE**	***z***	***pr* (>|*z*|)**
Low	Low	N – Y	2.55	2.94	0.87	0.385
	High	N – Y	−14.02	4.96	−2.82**	0.005
High	Low	N – Y	4.47	3.79	1.18	0.239
	High	N – Y	21.04	7.02	3.00**	0.003
Moderator levels						
Phonological similarity	**Context**	**Contrast**	**Estimate**	**SE**	***z***	***pr* (>|*z*|)**
Low	N	High – Low	1.93	1.45	1.33	0.183
	Y	High – Low	10.22	2.84	3.60***	< 0.001
High	N	High – Low	4.45	2.22	2.00*	0.045
	Y	High – Low	−3.84	1.60	−2.40*	0.016
Moderator Levels						
Familiarity	**Context**	**Contrast**	**Estimate**	**SE**	***z***	***pr* (>|*z*|)**
Low	N	High – Low	6.43	1.87	3.43***	< 0.001
	Y	High – Low	5.47	1.47	3.73***	0.002
High	N	High – Low	8.95	3.94	2.27*	0.023
	Y	High – Low	−8.59	2.14	−4.02***	< 0.001

*$p<0.05$, **$p<0.01$, ***$p<0.001$.
N, isolated condition. Y, contextual condition.
SE, standard error; *df*, degree of freedom.
High, Mean + *1SD*. Low, Mean-*1SD*.

TABLE 6 Results of LME model analysis of reaction times.

Variables	Estimate	SE	*df*	*t*	*pr* (>|*t*|)
Intercept	3.16	0.02	43.41	167.92***	<0.001
Familiarity	−0.06	0.01	45.44	−6.20***	<0.001
English Vocabulary Proficiency	0.00	0.02	29.01	0.23	0.816
ContextY	−0.03	0.02	41.07	−1.50	0.141
English Vocabulary Proficiency: ContextY	0.02	0.01	1108.40	2.00*	0.046

*$p < 0.05$, ***$p < 0.001$.
Participants = 31. Items = 44. Total observation = 1181.
SE, standard error. *df*, degree of freedom.
The optimal model is *lmer* [logrt ~ Vocabulary Proficiency + Context + Familiarity + Vocabulary Proficiency: Familiarity + (1|item) + (1|participant), data = data, lmerControl optimizer = "bobyqa," optCtrl = list (maxfun = 100000)].

TABLE 7 Results of simple main effects of context and English vocabulary proficiency.

Moderator Levels	Contrast	Estimate	SE	*df*	*t*	*pr* (>|*t*|)
N	High – Low	−0.00	0.02	32.17	−0.22	0.831
Y	High – Low	0.01	0.02	32.05	0.67	0.506
Low	N – Y	0.04	0.02	55.20	2.14*	0.037
High	N – Y	0.01	0.02	55.41	0.65	0.519

*$p < 0.05$.
N, isolated condition. Y, contextual condition.
SE, standard error; *df*, degree of freedom.
High, Mean + *1SD*. Low, Mean-*1SD*.

5. Discussion

While several studies on the processing of Japanese loanwords exist (e.g.,Yamato and Tamaoka, 2013; Tamaoka, 2018), they have often focused only on a limited range of variables. Moreover, there has been a lack of systematic exploration of the processing of loanwords by advanced Chinese JFL learners. Consequently, this study investigates the influence of English-Japanese phonological similarity, familiarity, context, and English vocabulary proficiency on the processing of Japanese loanwords in advanced Chinese JFL learners. Furthermore, this study develops a loanword processing model specifically tailored to advanced Chinese JFL learners. The results will serve as a useful reference for the acquisition of Japanese loanwords.

5.1. The influencing factors of loanword processing

Based on accuracy rates, phonological similarity had a significant influence on loanword processing, enhancing a learner's ability to accurately understand loanwords. However, we found no significant impact on reaction time in contrast with previous research results (Yamato and Tamaoka, 2013; Tamaoka, 2018). It was formulated the hypothesis that the phonological similarity would likely be facilitating the processing of Japanese loanwords. Consequently, the aforementioned results partly align with Hypothesis 1. Here, we discuss why no significant impacts existed in terms of reaction time. Advanced learners, due to their high level of Japanese proficiency, rely less on the phonetic representation of English than intermediate learners. Perhaps they have reached a stage where Japanese lexical processing is highly automated, enabling them to access conceptual representations directly without relying too much on English as an intermediary step. To fully understand how learner reliance on L2 and L3 processing evolves during different stages of acquisition, future research should include discussions on elementary and intermediate learners.

Regarding familiarity, in terms of both accuracy rates and reaction times, our findings indicate significant effects of familiarity on loanword processing. Higher levels of familiarity with loanwords were associated with higher accuracy rates and shorter reaction times. This suggests that familiarity plays a crucial role in loanword processing, even for advanced learners. These results, supporting Hypothesis 2, align with previous studies (Yamato et al., 2010; Ang et al., 2016; Li et al., 2020) and emphasize the stable facilitating effect of familiarity on loanword processing across different language pairs.

In terms of context, despite the lack of significant main effects, interactions between context and other factors existed. The results of the simple main effects analysis indicate that the presence of context did not always facilitate the processing of loanwords, thus providing insufficient evidence to support Hypothesis 3. This finding is inconsistent with previous results (e.g., Sereno et al., 2003; Goldwater et al., 2009). An interaction between context and phonological similarity was observed in terms of accuracy rates. This indicates that, when phonological similarity was low, context served as a compensatory and facilitating factor. Nevertheless, when phonological similarity was high, the presence of context resulted in an inhibitory effect. This may have been because, under conditions of high phonological similarity, learners partially relied on their L2 English vocabulary to process the loanwords. Therefore, excessive contextual information can increase learner language processing load, leading to an inhibitory effect.

The analysis of the second-order interaction revealed that the inhibitory effect of context was strongest when both phonological similarity and familiarity were high. This suggests that, when both phonological similarity and familiarity are high, perhaps learner processing of loanwords reaches a high level of automatization. Under such circumstances, the presence of context may introduce a certain degree of interference in learner processing, thereby reducing accuracy rates. Additionally, when both phonological similarity and familiarity are low, although the accuracy rate under the contextual condition is relatively low, it does not reach a significant level. Finally, the analysis of reaction times indicates that when learner English vocabulary proficiency is low, context facilitates the rapid processing of loanwords. However, for learners with high vocabulary proficiency, the facilitating effect of context is not significant. These results indicate that the impact of context on loanword processing depends on the specific characteristics of the loanwords (i.e., English-Japanese phonological similarity and familiarity with loanwords) and learner language proficiency (i.e., the proficiency of English and Japanese). The interaction between context and other factors in the processing of loanwords in advanced Chinese JFL learners will be further discussed in Section 5.2.

Lastly, there were no significant differences in English vocabulary proficiency in terms of accuracy rates and reaction times, which does not support Hypothesis 4. This finding is inconsistent with the results of previous studies (Tamaoka, 1997; Yamato and Tamaoka, 2013). There are two possible reasons for this result: (1) all participants in this study were advanced learners of Japanese, and the self-evaluation results of participant language proficiency indicated that their L3 Japanese proficiency was significantly higher than their L2 English proficiency. The frequency of L3 Japanese use was also significantly higher than that of L2 English use. This suggests that, for advanced learners, the impact of L2 English proficiency on the processing of loanwords in L3 Japanese among Chinese JFL learners is limited. (2) Although the influence of English vocabulary proficiency was not significant, as mentioned earlier, the influence of phonological similarity between Japanese and English still existed. We speculate that this might be related to the methodology used for testing English vocabulary proficiency in this study. The assessment of English vocabulary proficiency in this study employed a visual presentation method. Further investigation is needed to examine whether the processing of L3 Japanese loanwords among advanced Chinese JFL learners would be influenced by English vocabulary proficiency if an auditory presentation method were used.

5.2. The model of loanword processing in advanced Chinese JFL learners

As mentioned earlier, there is a lack of study on the processing model of Japanese loanwords by Chinese JFL learners, especially in terms of multiple perspectives and factors. Therefore, in this section, a processing model for loanwords used by advanced Chinese JFL learners is constructed. The preceding discussion shows that familiarity has the most significant impact among the four factors examined. In contrast, the effects of other factors are often limited and interconnected, exerting themselves collectively. Therefore, we propose that familiarity with Japanese loanwords should be considered the primary factor in loanword processing. Contextual information then follows, whereas phonological similarity and English vocabulary proficiency have a minor influence. To construct a processing model for advanced Chinese JFL learners, we developed a revised hierarchical model (Figure 3) based on previous studies (Kroll and Steward, 1994) and the aforementioned analysis.

According to the self-evaluation results for participant language proficiency and the lexical processing model used in previous studies (Chen and Leung, 1989; Kroll and Steward, 1994), learners can be characterized as unbalanced trilinguals, with Chinese, Japanese, and English ranked from highest to lowest in terms of proficiency. In Figure 3, lines ①–⑥ represent the links between a lexical representation and conceptual representation of Chinese, Japanese, and English. The thickness of a line indicates the strength of a link. During processing in advanced Chinese JFL learners for the acquisition of L3 Japanese, the paths through the lexical link (lines ②→①) or the conceptual link (line ③) are the most significant. On the contrary, the processing path through the L2 English lexical representation (lines ⑤→⑥) may not be the dominant pathway. Additionally, a processing path via Chinese and English (lines ②→④→⑥) is unlikely to exist. Research on the English lexical processing of Chinese English learners indicates that, as learner proficiency improves, the lexical processing mechanism transitions from a lexical link to a conceptual link (Chen and Leung, 1989). Considering the overwhelming influence of familiarity and the relatively weak correlation between Japanese loanwords and Chinese lexical representation, it is believed that, from an overall perspective, advanced Chinese JFL learners tend to directly access L2 Japanese loanwords using conceptual representation (line ③). However, a link is also influenced by other factors, such as phonological similarity, context, and English vocabulary proficiency.

The analysis of accuracy rates reveals that when there is low phonological similarity and high familiarity, context promotes learners' correct judgments. This means that, due to the presence of context, learners activate the conceptual representation in advance with the assistance of context. As a result of this activation, the three lexical representations of Chinese (line ①), Japanese (line ③), and English (line ⑥) are also activated. Since the experimental stimuli are in Japanese with low phonological similarity to English, the activation of English lexical representations should be the weakest (i.e., line ⑥'). However, when both phonological similarity and familiarity are high, the context reduces the accuracy rate. This indicates that with the activation of English lexical representation (line ⑥), due to the high phonological similarity, where the English lexical representation that was originally not involved in the bilingual competition, is now included in the competition, resulting in the emergence of trilingual competition. In this kind of situation, if learners are unable to

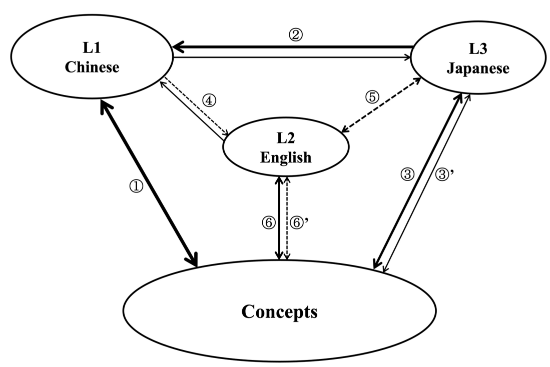


FIGURE 3
Processing of Japanese loanwords by advanced Chinese JFL learners under isolated and contextual conditions.

effectively utilize the activated English representations, their processing accuracy decreases. On the other hand, the reason for the disappearance of a significant effect from the context when both phonological similarity and familiarity are low, may be because that now it is highly challenging for learners, even with the activation of conceptual representation facilitated by the presence of context, to give correct responses. Moreover, although the context does not now exert a significant impact on the accuracy rate, the accuracy rate under the contextual condition is relatively low. This could be due to the increased cognitive load imposed on learners when processing contextual information, leading to a decrease in accuracy rate. In summary, when both phonological similarity and familiarity are high, the presence of context inhibits performance due to trilingual competition. In contrast, when both phonological similarity and familiarity are low, the lack of significant difference with context is possibly due to the additional cognitive load imposed by the context on learners in an already challenging situation (i.e., low familiarity and low phonological similarity), which results in a negative impact.

The results for reaction time show that the interaction between context and English vocabulary proficiency is significant. Context has a significant facilitating effect on learners with low proficiency but not on learners with high proficiency. Therefore, in the absence of context, the processing path for learners with low proficiency is represented by line ③', but with context, it transitions from line ③' to line ③, gradually forming a direct connection in a rapid response model. Therefore, context has the potential to assist in rapid judgment when English proficiency is low. However, if there is significant competition between the three languages (Chinese, English, and Japanese), it can result in a decrease in processing efficiency. In such a case, the presence of context alone may not be sufficient to overcome the challenges posed by language competition to ensure accurate processing. Therefore, while context can be beneficial in certain scenarios, it is important to mitigate potential interference from the simultaneous activation of multiple languages to maintain processing accuracy and efficiency.

5.3. Suggestions on Chinese JFL learners' acquisition of Japanese loanwords

Familiarity significantly influences the processing of loanwords. Previous studies on Japanese language education have indicated that Chinese JFL learners, especially beginners or intermediate learners, tend to avoid using loanwords (Deng, 2018). This implies that Chinese JFL learners, influenced by their L1, have a tendency to prefer using *Kanji* words over loanwords when encountering Japanese lexicons with the same meaning that include both options. The results of the current study revealed that even advanced JFL learners exhibited an average reaction time of 1548.22 ms for all loanwords, indicating room for improvement in their response speed. Hence, it is crucial for learners to actively engage with loanwords, gradually enhancing their familiarity with them.

The results indicate that English-Japanese phonological similarity and English vocabulary proficiency have a weak influence on the processing of loanwords in advanced JFL learners. This suggests that the link between Japanese lexical representation and English lexical representation is weak for Chinese learners. If learners cannot handle the bilingual competition between English and Japanese well, it decreases the accuracy and speed with which they process Japanese loanwords. Therefore, to maximize the benefits of learners' existing English knowledge, it is recommended that they focus on word-pair learning between English and Japanese when teaching and learning loanwords. Because of the competition between the three languages, it is advised that paired learning between English and Japanese be used. This approach enhances bilingual coordination and mitigates the negative impact of competition.

We observed that context does not always have a positive impact on advanced Chinese JFL learners and can even have an inhibitory effect. Therefore, when teaching Japanese, it is crucial to carefully consider the characteristics of the context and select appropriate materials that facilitate learner acquisition. Based on the findings of this study, it is recommended that teachers analyze the specific characteristics of loanwords, make thoughtful adjustments to the contextual materials, and implement targeted teaching strategies to cater to the needs of advanced Chinese JFL learners.

6. Conclusion

In this study, a lexical decision task was utilized to examine how English-Japanese phonological similarity, familiarity, context, and English vocabulary proficiency impacted the processing of Japanese loanwords among Chinese JFL learners. An analysis using a (generalized) linear mixed-effect model showed that the influence of English-Japanese phonological similarity, English vocabulary proficiency, and context on Japanese loanword processing was not always positive. As learners' Japanese proficiency improved, they tended to process Japanese loanwords directly based on conceptual representations, with English vocabulary proficiency showing no significant influence. When both Japanese familiarity and English-Japanese phonological similarity were high, context exerted an inhibitory effect. These conclusions underscored the complexity of examining the lexical processing mechanism in trilingual individuals and emphasized the need to consider various influencing factors when investigating the associations between representations.

This study has several limitations. The participants consisted solely of advanced Chinese JFL learners, which may reduce the generalizability of the findings when considering Chinese JFL learners at all proficiency levels. Consequently, it would be valuable to expand the scope of participants to include elementary and intermediate Chinese JFL learners and compare the results with the findings of the current study, to further investigate how the processing of Japanese loanwords by Chinese JFL learners changes as their Japanese proficiency improves. Additionally, as a special component of Japanese vocabulary, to more deeply understand how these factors examined in this study play their roles, further exploration of the processing of Japanese loanwords by English JFL learners and Japanese EFL learners would present an intriguing avenue for further investigation. By comparing the findings of such a study with the current research, we can gain additional insights into the processing mechanism of Japanese loanwords.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving humans were approved by the Beijing Center for Japanese Studies, Beijing Foreign Studies University. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

YG, QS, and XF contributed to the conception of the work and revised the manuscript and confirmed its final version. YG and XF collected the experimental data. QS conducted the data analysis. YG and QS wrote the first manuscript. All authors contributed to the article and approved the submitted version.

Funding

This study was supported by the "Excellent Talents Program of Beijing Foreign Studies University 2021 (2021-ZYQNJS-011).

Conflict of interest

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

Publisher's note

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2023.1224830/full#supplementary-material

References

Ang, C. (昂晨), Lv, H. (吕欢), Zhou, Y. (周亚聪), Li, B. (李博闻), and Wang, R. (王瑞明) (2016). The familiarity influence on the activation of non-target language in language comprehension of unskilled Chinese-English bilinguals (词汇熟悉度对非熟练中英双语者语言理解转换中非目标语言激活的影响). *Psychol. Dev. Educ. (*心理发展与教育*).* 32, 26–32. doi: 10.16187/j.cnki.issn1001-4918.2016.01.04

Antón, E., and Duñabeitia, J. A. (2020). Better to be alone than in bad company: cognate synonyms impair word learning. *Behav. Sci.* 10:123. doi: 10.3390/bs10080123

Balota, D. A., Alexander, P., and Keith, R. (1985). The interaction of contextual constraints and parafoveal visual information in reading. *Cogn. Psychol.* 17, 364–390. doi: 10.1016/0010-0285(85)90013-1

Bates, D., Maechler, M., Bolker, B., and Walker, S. (2015). Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* 67, 1–48. doi: 10.18637/jss.v067.i01

Chen, H. C., and Leung, Y. S. (1989). Patterns of lexical processing in a nonnative language. *J. Exp. Psychol. Learn. Mem. Cogn.* 15, 316–325. doi: 10.1037/0278-7393.15.2.316

Crocetta, T. B., and Andrade, A. (2015). The problem of measuring reaction time using software and hardware: a systematic review (Retrasos en la medición del tiempo con el uso de computadoras en la investigación del Tiempo de Reacción: Una revisión sistemática). *Revista De Psicologia Del Deporte.* 24, 341–349.

De Angelis, G., and Selinker, L. (2001). "Interlanguage transfer and competing linguistic systems in the multilingual mind," in *Cross-Linguistic Influence in Third Language Acquisition: Psycholinguistic Perspectives*. eds. J. Cenoz, B. Hufeisen and U. Jessner (Clevedon, UK: Multilingual Matters), 42–58.

De Groot, A. M., Delmaar, P., and Lupker, S. J. (2000). The processing of interlexical homographs in translation recognition and lexical decision: support for non-selective access to bilingual memory. *Q. J. Exp. Psychol. A.* 53, 397–428. doi: 10.1080/713755891

Deng, Q. (鄧琪) (2018). Use of Katakana-loanwords by Chinese learners: an analysis based on I-JAS (中国人日本語学習者の外来語使用に対する一考察―「多言語母語の日本語学習者横断コーパス」を用いた調査をふまえて―). *Learn. Corpus Stud. Asia World.* 3, 241–261. doi: 10.24546/81010130

Dewaele, J. M. (1998). Lexical inventions: French interlanguage as L2 versus L3. *Appl. Linguis.* 19, 471–490. doi: 10.1093/applin/19.4.471

Dijkstra, T., and Walter, J. V. H. (2018). Visual word recognition in multilinguals. *Oxford Handbook Psycholinguist.* 2, 118–143.

Fei, X. (費暁東), and Li, H. (李海鹏) (2017). Effects of sentence constraint on processing of auditorily presented words in Chinese intermediate learners of the Japanese language: an experimental study with manipulation of orthographical and phonological similarities between Chinese and Japanese Kanji characters (文の制約性が中国人中級日本語学習者の聴覚的単語認知過程に及ぼす影響―中日漢字の形態・音韻類似性を操作した実験的検討―). *Theory Res. Dev. Learn. Syst. (*学習システム研究*).* 5, 29–44. doi: 10.15027/42659

Fei, X., Zhao, S., and Liu, J. (2022). Auditory recognition of Chinese-Japanese cognates and homographs by Chinese JFL learners. *Psychologia* 64, 1–22. doi: 10.2117/psysoc.2021-A144

Ferré, P., Sánchez-Casas, R., and Guasch, M. (2006). Can a horse be a donkey? Semantic and form interference effects in translation recognition in early and late proficient and nonproficient Spanish-Catalan bilinguals. *Lang. Learn.* 56, 571–608. doi: 10.1111/j.1467-9922.2006.00389.x

Geng, Y. (耿耀耀) (2022). The influencing factors of loanwords acquisition of Chinese Japanese learners: an empirical analysis based on the modified grounded theory approach (中国語を母語とする日本語学習者の外来語習得に関する研究―M-GTAを用いた質的検討―). *Comparative Japanese Studies Annual Bulletin (*比較日本学教育研究部門研究年報*).* 18, 68–74.

Goldwater, S., Griffiths, T. L., and Johnson, M. (2009). A Bayesian framework for word segmentation: exploring the effects of context. *Cognition* 112, 21–54. doi: 10.1016/j.cognition.2009.03.008

Hermans, D., Bongaerts, T., De Bot, K., and Schreuder, R. (1998). Producing words in a foreign language: can speakers prevent interference from their first language? *Biling. Lang. Congn.* 1, 213–229. doi: 10.1017/S1366728998000364

Hoshino, N., and Kroll, J. F. (2008). Cognate effects in picture naming: does cross-language activation survive a change of script? *Cognition* 106, 501–511. doi: 10.1016/j.cognition.2007.02.001

Jha, B., and Chang, X. (常笑), Lin, Y. (林韻), Wang, X. (王校偉), and Matsumi, N. (松見法男) (2018). Interpreting unknown words of advanced-level Chinese learners of Japanese: experimental analysis using manipulation of word notation and sentence constraints (中国語を母語とする上級日本語学習者の未知語の意味推測過程―単語表記と文の制約性を操作した実験的検討―). *Bulletin of the graduate School of Education, Hiroshima University. II, Arts and science education (*広島大学大学院教育学研究科紀要.第二部, 文化教育開発関連領域*)* 67, 173–180. doi: 10.15027/46783

jamovi project. (2022). *Jamovi. (version 2.3) [computer software]*. Available at: https://www.jamovi.org.2023.03.28 access.

Jankowiak, K. (2021). Current trends in electrophysiological research on bilingual language processing. *Lang. Linguist. Compass* 15:e12436. doi: 10.1111/lnc3.12436

Jasone, C. (2001). "The effect of linguistic distance, L2 status and age on cross-linguistic influence in the third language acquisition" in *Cross-linguistic influence in third language acquisition: psycholinguistic perspectives, Clevedon* (England: Multilingual Matters), 8–20.

Kess, J. F., and Miyamoto, T. (2000). "A history of the Japanese orthography" in *The Japanese mental lexicon: psycholinguistic studies of Kana and Kanji processing* (Amsterdam, The Netherlands: John Benjamins Publishing), 13–32.

Kroll, J. F., and Steward, E. (1994). Category interference in translation and picture naming: evidence for asymmetric connections between bilingual memory representations. *J. Mem. Lang.* 33, 149–174. doi: 10.1006/jmla.1994.1008

Kroll, J. F., Van Hell, J. G., Tokowicz, N., and Green, D. W. (2010). The revised hierarchical model: a critical review and assessment. *Biling. Lang. Congn.* 13, 373–381. doi: 10.1017/S136672891000009X

Kuznetsova, A., Brockhoff, P. B., and Christensen, R. H. B. (2017). lmerTest package: tests in linear mixed effects models. *J. Stat. Softw.* 82, 1–26. doi: 10.18637/jss.v082.i13

Lee, Y., Jang, E., and Choi, W. (2018). L2-L1 translation priming effects in a lexical decision task: evidence from low proficient Korean-English bilinguals. *Front. Psychol.* 9:267. doi: 10.3389/fpsyg.2018.00267

Li, M., and Kirby, J. R. (2015). The effects of vocabulary breadth and depth on English Reading. *Appl. Linguis.* 36, amu007–amu634. doi: 10.1093/applin/amu007

Li, L. (李利), Mo, L. (莫雷), and Wang, R. (王瑞明) (2011). Regulation of the familiarity of second language words in bilinguals' semantic access (二语词汇熟悉度在双语者语义通达中的调节作用). *J. Psychol. Sci. (*心理科学*).* 34, 799–805. doi: 10.16719/j.cnki.1671-6981.2011.04.013

Li, P., Zhang, F., Yu, A., and Zhao, X. (2020). Language history questionnaire (LHQ3): an enhanced tool for assessing multilingual experience. *Biling. Lang. Congn.* 23, 938–944. doi: 10.1017/S1366728918001153

Matsumi, N. (松見法男), Fei, X. (費暁東), and Cai, F. (蔡鳳香) (2012). "The lexical processing of Japanese Kanji words (日本語漢字単語の処理過程—中国語を母語とする中級日本語学習者を対象とした実験的検討—)" in *Second language acquisition research and language education (*第二言語習得研究と言語教育*)*, eds K. Hatasa, Y. Hatasa, M. Kudara and T. Shimizu Tokyo: Kuroshio Shuppan Press, 43–67. [in Japanese]

Mochizuki, T. (望月通子) (2012). Learning and teaching material development for Katakana loanwords: feedback and design (基本語化を考慮したカタカナ外来語の学習と教材開発—その振り返りと新たな開発に向けて—). Journal of foreign language studies. *Kansai Univ. Faculty Foreign Lang Bulletin (*関西大学外国語学部紀要*).* 6, 1–16.

Mori, Y. (2014). Review of recent research on Kanji processing, learning, and instruction. *Japanese Lang Lit.* 48, 403–430.

Nation, I. S. P. (1983). Testing and teaching vocabulary. *Guidelines* 5, 12–25.

Nation, I. S. P., and Beglar, D. (2007). A vocabulary size test. *Lang. Teach.* 31, 9–13.

Peng, G., and Moriya, M. (2007). *Comprehensive Japanese (*综合日语*).* Beijing: Peking University Press.

Perea, M., Soares, A. P., and Comesaña, M. (2013). Contextual diversity is a main determinant of word identification times in young readers. *J. Exp. Child Psychol.* 116, 37–44. doi: 10.1016/j.jecp.2012.10.014

R Core Team. (2022). *R: a language and environment for statistical computing.* Vienna: R Foundation for Statistical Computing.

Sereno, S. C., Brewer, C. C., and O'Donnell, P. J. (2003). Context effects in word recognition: evidence for early interactive processing. *Psychol. Sci.* 14, 328–333. doi: 10.1111/1467-9280.14471

Silverberg, S., and Samuel, A. G. (2004). The effect of age of second language acquisition on the representation and processing of second language words. *J. Mem. Lang.* 51, 381–398. doi: 10.1016/j.jml.2004.05.003

Singh, N., Mishra, B. B., Bajpai, S., Singh, R. K., and Tiwari, V. K. (2014). Natural product based leads to fight against leishmaniasis. *Bioorganic Med. Chem.* 22, 18–45. doi: 10.1016/j.bmc.2013.11.048

Song, Q. (宋启超), and Fei, X. (费晓东) (2022). The effect of semantic transparency, translational congruency, and context on L2 Japanese collocational processing (语义透明度、同译性及语境对日语二语语块加工的影响). *Mod. Foreign Lang. (*现代外语*).* 45, 659–670.

Song, Q., Fei, X., and Matsumi, N. (2023). The lexical processing of Japanese collocations by Chinese Japanese-as-a-foreign-language learners: an experimental study by manipulating the presentation modality, semantic transparency, and translational congruency. *Front. Psychol.* 14:1142411. doi: 10.3389/fpsyg.2023.1142411

Stanovich, K. E., and Richard, F. W. (1983). On priming by a sentence context. *J. Exp. Psychol. Gen.* 112, 1–36. doi: 10.1037/0096-3445.112.1.1

Sube, M. (須部宗生) (2012). Katakana English and Japanese English: Focusing on their Recent Trends (カタカナ英語と和製英語-最近の傾向を中心として-). *Environment and management: Journal of Shizuoka Sangyo University* (環境と経営: 静岡産業大学論集). 19, 127–137.

Tajima, K., Erickson, D., and Nagao, K. (2002). Production of syllable structure in a second language: factors affecting vowel epenthesis in Japanese-accented English. *IULC Work. Papers* 2, 77–91.

Tamaoka, K. (玉岡賀津雄) (1997). The processing strategy of words presented in Kanji and Kana by Chinese and English speakers learning Japanese (中国語と英語を母語とする日本語学習者の漢字および仮名表記語彙の処理方略). *Stud. Lang. Lit. (*言語文化研究*).* 17, 65–77.

Tamaoka, K. (玉岡賀津雄) (2018). Lexical connections among three languages: the effects of L1 Chinese and L2 English on the processing of L3 Japanese loanwords by native Chinese speakers learning Japanese (三言語間の語彙的結合—中国人日本語学習者によるL3日本語の外来語処理におけるL1中国語と第二言語英語の影響—). *Res. Japanese Lang. Educ. Chinese Speakers (*中国語話者のための日本語教育研究*).* 9, 17–34.

Van Hell, J. G., and Dijkstra, T. (2002). Foreign language knowledge can influence native language performance in exclusively native contexts. *Psychon. Bull. Rev.* 9, 780–789. doi: 10.3758/BF03196335

Wesche, M., and Paribakht, T. S. (1996). Assessing second language vocabulary knowledge: depth versus breadth. *Can. Mod. Lang. Rev.* 53, 13–40. doi: 10.3138/cmlr.53.1.13

Yamato, Y. (大和祐子), and Tamaoka, K. (玉岡賀津雄) (2011). The on-line processing of Kanji-and Katakana-presented words in Japanese texts: a comparison of greater and lesser lexical knowledge groups of native Chinese speakers learning Japanese (日本語テキストのオンライン読みにおける漢字表記語と片仮名表記語の処理—中国人日本語学習者の語彙能力上位群と下位群の比較—).*Papers of the Japanese Language Teaching Association in honor of Professor Fumiko Koide (*小出記念日本語教育研究会*)*, 19, 73–86. doi: 10.18993/jcrdajp.36.1_33

Yamato, Y. (大和祐子), and Tamaoka, K. (玉岡賀津雄) (2013). Effects of English knowledge via Japanese loanwords on reading of Japanese texts performed by native Chinese speakers learning Japanese (中国人日本語学習者による外来語処理への英語レキシコンの影響). *Lexicon Forum (*レキシコンフォーラム*).* 6, 229–267.

Yamato, Y. (大和祐子), Tamaoka, K. (玉岡賀津雄), and Chu, X. (初相娟) (2010). Effect of Japanese learning-length on the processing of loanwords and Kanji compounds by native Chinese speaking students (中国人日本語学習者による外来語および漢字語の処理における学習期間の影響). *Studia Linguistica* (ことばの科学).23, 101–119. doi: 10.18999/stul.23.101

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 22 September 2023
DOI 10.3389/fpsyg.2023.1266921

Check for updates

OPEN ACCESS

EDITED BY
Katarzyna Jankowiak,
Adam Mickiewicz University, Poland

REVIEWED BY
Katharina Oster,
Johannes Gutenberg University Mainz, Germany
Zhilong Xie,
Jiangxi Normal University, China

*CORRESPONDENCE
Ruiyuan Wang
✉ ruiyuanwang@sina.com

RECEIVED 25 July 2023
ACCEPTED 06 September 2023
PUBLISHED 22 September 2023

CITATION
Wang R, Han J, Di Biase B and Antoniou M (2023) The effect of study-abroad experience on lexical translation among interpreting students.
Front. Psychol. 14:1266921.
doi: 10.3389/fpsyg.2023.1266921

The effect of study-abroad experience on lexical translation among interpreting students

Ruiyuan Wang[1]*, Jing Han[2], Bruno Di Biase[2] and Mark Antoniou[3]

[1]School of International Studies, Shaanxi Normal University, Xi'an, Shaanxi, China, [2]School of Humanities and Communication Arts, Western Sydney University, Sydney, NSW, Australia, [3]The MARCS Institute for Brain, Behaviour, and Development, Western Sydney University, Sydney, NSW, Australia

This study investigates the impact of study-abroad experience (SAE) on lexical translation among 50 Chinese (L1)-English (L2) interpreting students. Participants were divided into two groups based on their experience abroad. Both groups consisted of 25 unbalanced L2 learners who were matched in age, working memory, length of interpreting training, and L2 proficiency. Bidirectional word translation recognition tasks, from L1 to L2 and L2 to L1, highlighted several key findings: (1) both groups were significantly more accurate and faster from L2 to L1 than in the reverse direction; (2) the study abroad (SA) group was more inclined to respond quickly at the risk of making errors, whereas the non-study abroad (NSA) group tended to be more cautious, prioritising accuracy over speed; (3) the SA group were more balanced and consistent in their performance across lexical translations in both directions than the NSA group. These results emphasise the potent effect of SAE in resolving bilinguals' language competition, especially in streamlining language switching, a cognitive process critical for interpreting students engaging daily with dual languages.

KEYWORDS

study-abroad experience, Chinese-English, lexical processing, word translation, interpreting students, direction-dependent asymmetry

1. Introduction

Interpreting, a linguistically complex and cognitively demanding activity, necessitates quick and accurate alternation between two languages within tight temporal constraints (Christoffels et al., 2003). In any interpreting mode, be it simultaneous or consecutive, a lapse or delay in language processing can potentially escalate the cognitive load on interpreters, consequently straining their working memory and affecting their overall interpreting performance (O'Brien et al., 2006). The intricacies of this task underline the paramount importance of efficient lexical retrieval and translation (Mead, 2002; Gile, 2009). Indeed, studies consistently highlight a positive correlation between the speed and accuracy of lexical translation and broader interpreting performance, attesting to the role of lexical processing in interpreting practice (Christoffels et al., 2003; Santilli et al., 2019).

Though some interpreters are often recommended to interpret solely into their L1, many possess the capability for bidirectional interpreting—comprehending in one language and interpreting into another. Yet, it is commonly observed that they may not perform equally well in both directions (Russell and Takeda, 2015). Bilingual individuals often display direction-dependent asymmetry in their lexical processing, with a faster and more accurate performance from their L2 to their L1 than in the reverse direction, indicating an advantage in this direction (Kroll and Stewart, 1994; De Groot and Poot, 1997; Green, 1998; Issa and Shyamala, 2021).

Increasing attention has been paid to factors influencing bilingual lexical translation. Evidence points towards the impact of variables such as participants' working memory (Sunderman and Kroll, 2009), L2 proficiency (Meuter and Allport, 1999; Costa, 2005; Berghoff et al., 2021; Issa et al., 2022), language use frequency (Christoffels et al., 2007), and language exposure (Kroll et al., 1998; Kroll and Sunderman, 2003; Linck et al., 2008; Kleinman et al., 2022). Among these, the study-abroad experience (SAE) holds significant implications.

SAE, within the field of second language acquisition, is characterised as a type of L2 learning setting that differs from both purely natural exposure and classroom instruction. While natural exposure pertains to the spontaneous, untutored acquisition of a language in its native country, classroom instruction often refers to the teaching of a foreign language in a country where that language is not the primary mode of communication. For instance, one might learn English in Chinese classrooms, with the classroom being the primary, if not the sole, exposure to the language (Muñoz, 2008). SAE, however, integrates formal classroom training with daily life experiences in a country where the target language is dominant, often after students have initially studied the language in their home countries (Xie and Dong, 2021).

Practically, this environment has been shown to increase L2 processing (Antoniou et al., 2015), suppress L1 dominance (Baus et al., 2013), enhance individual cognitive performance (Xie and Dong, 2021), and thus may make it generally easier for bilinguals to access and switch between the two languages (Bonfieni et al., 2019). Theoretically, SAE provides a unique context to examine and challenge existing bilingualism models. It could shed light on the complex interplay of exposure, cognition, and language utilisation in shaping bilingual lexical processing and help refine our understanding of bilingual language control mechanisms.

Most previous research on lexical development in the SAE context mainly concentrates on vocabulary knowledge growth, typically evaluated through word association and word recognition tasks. However, fewer studies have explored how SAE influences a bilingual's command of two languages, particularly in terms of efficiency (speed and accuracy) and asymmetry in bidirectional lexical translation. Importantly, interpreting students, who consistently keep both languages active even in predominantly monolingual environments, are a unique subgroup of bilinguals (Grosjean, 1997; Babcock and Vallesi, 2017). These students are in a distinct situation as they are training to be professional interpreters, a role that necessitates frequent and skilled language switching. Yet, despite this distinct situation and the demanding requirements of their future profession, few studies specifically investigate the impact of SAE on this group.

Investigating the impact of SAE on interpreting students' lexical translation performance could offer valuable insights into bilinguals' lexical processing. Additionally, it may guide interpreting educators in developing effective training tasks to meet specific pedagogical objectives in both study-abroad and home-country classroom settings.

2. Literature review

2.1. Theoretical framework

Various models have been developed to elucidate the intricacies of bilingual lexical processing. For instance, the Hierarchical Model by Kroll and Stewart (1994) delves into the interrelationships between L1 and L2 words and concepts. In contrast, other models emphasise the lexicon itself (for instance, the role of language nodes in the original BIA model) or advocate for an external mechanism to regulate the lexical system's operations.

Our current study predominantly draws upon the Inhibitory Control model introduced by Green (1998). As illustrated in Figure 1. This model is premised on the widely-held linguistic belief that during, or even prior to, speech comprehension and production, elements such as sounds, forms, and concepts from a bilingual's languages are activated simultaneously in a non-language-specific manner, leading to competition for selection (Odlin, 2003, 2012; Jarvis and Pavlenko, 2008).

The Inhibitory Control model postulates that the selection of the target language at any given time involves the suppression of the non-target one. This suppression process is overseen by the Supervisory Attentional System (SAS), a superior cognitive control mechanism that intervenes with the language system as needed. One of the key insights derived from the model is the relationship between language proficiency and inhibition. The more proficient a bilingual is in a language, the more effortlessly they can suppress the non-target language, and conversely, lesser proficiency might entail greater cognitive effort in this suppression (Meuter and Allport, 1999; Meuter, 2009). Inherent to this bilingual ability is the concept of 'switching costs', which are cognitive tolls associated with toggling between languages. 'Direction-dependent asymmetry' epitomises these costs. Given that the dominant L1 generally has a heightened activation compared to the less dominant L2, more cognitive resources are required to inhibit L1, thus facilitating L2 production. In essence, bilinguals often exhibit a faster, more accurate performance when switching from their L2 to their L1 compared to the reverse direction (Kroll and Stewart, 1994; De Groot and Poot, 1997; Green, 1998; Issa and Shyamala, 2021). With this intricate interplay between

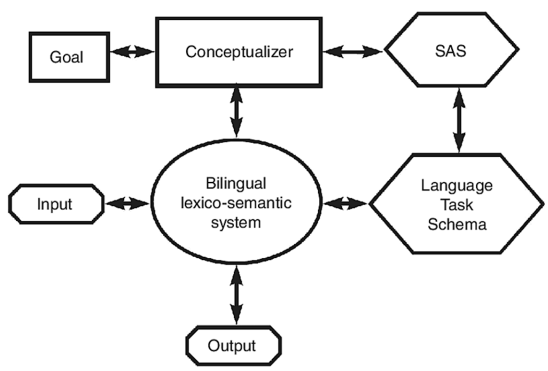


FIGURE 1
Diagram illustrating the Inhibitory Control model, adapted from Green (1998), showing the interplay of different cognitive processes in language translation.

bilingualism and cognitive functions in mind, it becomes imperative to explore how immersive bilingual experiences, such as SAE, further influence this dynamic.

2.2. Impact of study-abroad experience on bilinguals' cognitive performance

Lexical translation extends beyond a mere linguistic task; it is intrinsically woven with cognitive processes, especially those governed by working memory. As a fundamental pillar of cognitive resources, working memory is predictive of a host of intricate cognitive tasks, inclusive of language processing (Daneman and Merikle, 1996; Miyake and Friedman, 1998; Ardila, 2003). At its core, working memory boasts an executive function, acting as a cognitive controller (Baddeley, 1996). This controller is responsible for tasks such as selective attention, distraction inhibition, and overall coordination (Baddeley, 1996, 2017). Study-abroad experience, with its intensive L2 immersion, has the potential not just to mould lexical proficiency but also to influence the very cognitive mechanisms that underpin bilingual language processing (DeLuca et al., 2020).

Researchers suggest that the regular, habitual use of the bilingual control mechanism to reconcile the L1 and L2 competition should have cognitive benefits (Ransdell et al., 2006; Linck et al., 2008; Bialystok et al., 2009; Bartolotti et al., 2011; Xie and Dong, 2017). Owing to the constant regulation of two language systems as a result of SAE, bilinguals in this language environment are exposed to extensive practice of executive functions of language control on daily basis, which reduces the dominance of L1 and makes it generally easier to access and switch between the two languages (Bonfieni et al., 2019; Xie and Dong, 2021).

A substantial body of evidence indicates that bilingual cognitive control capabilities can be moulded by diverse bilingual experiences and language use frequency (e.g., Costa et al., 2009; Prior and Gollan, 2011; Xie and Dong, 2017; Xie and Dong, 2021). For instance, Xie and Dong (2017), provided empirical support for the assertion that public speaking training experience can enhance bilinguals' cognitive control capabilities. Furthering this line of research, Xie and Dong (2021) examined the potential cognitive control disparities between bilinguals with SAE and those without. Despite matching both groups based on demographic factors like age, socioeconomic status, and intelligence, the study found that SAE was associated with a significant advantage in cognitive control, particularly in mental set shifting. This enhancement was attributed to increased L2 usage and switching within the SAE context. As SAE seemingly enhances cognitive mechanisms through frequent L2 use, it's plausible to inquire how this immersive experience specifically impacts lexical processing, an integral part of bilingual language utilisation.

2.3. Impact of study-abroad experience on lexical processing

Despite the prevalence of SAE in L2 learning programmes, research examining the effects of SAE on lexical processing remains limited and somewhat inconclusive (Hamano-Bunce et al., 2019). The majority of prior research has focused on the impact of SAE on L2 receptive vocabulary knowledge growth (e.g., Fitzpatrick, 2012; Issa et al., 2020; Kleinman et al., 2022), or L2 productive vocabulary development (Segalowitz and Freed, 2004; Llanes and Muñoz, 2009; Barquin, 2012; Pérez-Vidal et al., 2012; Lara, 2014), or its contribution to enhanced sensitivity to L2 speech (Grey et al., 2015). However, a closer examination reveals that the cognitive advantages facilitated by SAE have profound implications on lexical processing.

For instance, within the realm of SAE and L2 lexical processing, Chinese speakers who spent an average of 6.5 years in the US demonstrated better performance in their speed and accuracy in recognising spoken L2 words. This suggests that SAE participants may possess a heightened capability to decipher talker variability and expedite nonnative language processing with fewer cognitive resources than their counterparts without such experience (Antoniou et al., 2015).

The avenue of research addressing SAE's influence on the decline of L1 availability has also yielded significant insights. Notably, studies have indicated that extended immersion in an L2 context amplifies the cognitive effort required for languages that are not regularly practised or used (Tu et al., 2015). A substantial body of research, often focusing on typologically similar language pairs like English-Spanish, has showcased a decline in L1 lexical representation during immersion in L2 settings (Linck et al., 2009; Kaushanskaya et al., 2011; Baus et al., 2013).

On another front, SAE also appears to shape bilinguals' communication behaviours. A study by Tokowicz et al. (2004) observed that individuals with SAE experience often ventured answers even when they were uncertain about accurate word translations, suggesting a bolstered propensity to communicate irrespective of potential inaccuracies.

Collectively, these studies highlight the intricate interplay between the cognitive mechanisms honed through SAE and their subsequent manifestation in bilingual lexical processing.

2.4. Research gaps and aims

Fundamental questions remain regarding the impact of SAE on lexical translation. Notably, studies on Chinese-English language pairs, particularly with Chinese as L1 and English as L2, are significantly underrepresented—this is significant given the rising trend of Chinese students studying abroad in English-speaking countries (Bhandari, 2017; Pavlacic, 2018; Szego, 2020). The unique linguistic characteristics of interpreting students due to their constant activation of both languages have been largely sidelined in SAE research. Such linguistic training may interact with SAE in ways that either mitigate or amplify its impact. Furthermore, the relationship between bidirectional lexical translation and SAE has been sparsely explored. Previous studies have often examined the impact of SAE on either L1 or L2 lexical processing, neglecting the potential influence of SAE on the bidirectional lexical translation efficiency and asymmetry.

In light of these identified gaps, our study seeks to determine the effect of SAE on bidirectional word translation tasks among interpreting students with Chinese as L1 and English as L2. Specifically, we examine how SAE impacts their efficiency (measured by accuracy and response times) and direction-dependent asymmetry in bidirectional translation.

Drawing from the literature that highlights the benefits of SAE, we hypothesise that participants with SAE in English-speaking countries will demonstrate better efficiency and balance in bidirectional word translation tasks compared to those without such experience.

3. Materials and methods

3.1. Participants

To ensure a homogenous sample, we only included participants who shared a similar linguistic and cultural background. All participants were native Chinese speakers, aged between 20 and 30, a criterion chosen based on existing research indicating the influence of age and cultural background on brain activity (Signorelli et al., 2012; Han and Ma, 2014).

As detailed in Table 1, the study comprised 50 participants, all of whom were pursuing master's degrees in Chinese-English bidirectional interpreting and translation. They were all in their third semesters and had commenced mandatory English education at the age of 12, following the establishment of their L1. Participants were evenly divided into two groups based on their SAE.

The study abroad (SA) group, consisted of 25 students, each enrolled in one of these three universities in Sydney, Australia (nine from Western Sydney University, eight from Macquarie University and eight from the University of New South Wales). They had an average SAE duration of 3.64 years and an average age of 25. In their Australian academic environment, they typically engaged in interpreting and translation classes conducted mainly in English for 25–30 h weekly. While interactions with faculty and international peers were predominantly in English, they reverted to Chinese for conversations with fellow Chinese students.

On the other hand, the non-study abroad (NSA) group was composed of 25 students from three universities in Xi'an, mainland China (seven from Shaanxi Normal University, eight from Xi'an Jiaotong University and 10 from Xi'an International Studies University). These students had not ventured to English-speaking nations and had an average age of 24. Their curriculum consisted of approximately 28–30 h of interpreting and translation weekly. However, the teaching methodology leaned heavily on their native Chinese language, and their daily interactions seldom involved English.

TABLE 1 Descriptive information of participant groups (mean and standard deviation).

Characteristic	SA group	NSA group
Number of participants	25.0	25.0
Age (years)	25.0 (0.31)	24.0 (1.19)
Years of study abroad	3.64 (0.46)	0 (0)
University semester of interpreting learning	3.0 (0)	3.0 (0)
Age of onset (learning L2)	12.1 (0.65)	12.1 (0.73)
L2 Interacting with English speakers	37.5** (2.04)	19.9** (1.40)
Watching TV	25.0 (0.96)	27.7 (1.87)
Reading	32.0 (1.36)	31.9 (1.91)
Self-instruction	26.4 (0.71)	26.4 (0.68)

*$p < 0.05$, **$p < 0.01$, and ***$p < 0.001$.

To further elucidate participants' linguistic profiles, we utilised the Language Experience and Proficiency Questionnaire (LEAP-Q). Widely recognised in linguistic and psycholinguistic research (e.g., Bialystok et al., 2009), this tool provided insights into participants' weekly L2 exposure. Notably, the SA group reported significantly higher weekly interaction hours with English speakers than the NSA group, though other linguistic activities remained comparable for both groups.

3.2. Materials

3.2.1. Working memory span

Participants' working memory resource availability was assessed using the English reading span task (Daneman and Carpenter, 1980), administered via E-Prime Professional 2.0. Sentences were displayed on the screen one at a time, and participants were instructed to read each sentence aloud, evaluate its semantic plausibility, and attempt to remember the final word of each sentence. Sentences were presented in sets of increasing size, ranging from two to five sentences. There were three series for each set size, resulting in a total of 42 sentences. Each sentence contained 11 to 13 words and concluded with a distinct word. Half of the sentences in this task were semantically plausible, while the remaining half were implausible.

Upon completion of each set, participants were prompted to recall as many sentence-final words as they could. No restrictions were imposed on the order or duration of the recall. Participants' recalled final words were considered valid only if accompanied by accurate judgements regarding sentence plausibility. The number of correctly recalled final words served as an indicator of a participant's working memory span.

3.2.2. L2 proficiency

This study emphasises L2 proficiency due to the background of our participants. Being native Chinese speakers, their proficiency in their L1 is uniformly high and consistent for daily communication. In contrast, as English is their foreign language, their L2 proficiency is expected to vary. They were asked to complete LexTALE, a reliable and standardised online test of general English proficiency that is widely employed in linguistic studies (e.g., Christoffels et al., 2007; Lemhöfer and Broersma, 2012; De Bruin et al., 2014; Keuleers et al., 2015). This assessment required participants to determine whether a given sequence of letters displayed on the screen constituted an English word. Upon completion of the task, the participants' scores were immediately calculated and displayed on the screen.

3.2.3. Bidirectional word translation

Participants' lexical processing performances were assessed using word translation recognition tasks, administered via E-Prime Professional 2.0 and Chronos. Distinct versions of the word translation recognition task were employed for both language directions (English-Chinese and Chinese-English). The stimuli consisted of 60 English and 60 Chinese words, which were presented auditorily in a randomised order for each participant using E-Prime Professional 2.0. During each trial, participants listened to a word through headphones, and the correct translation equivalent or an incorrect, misleading word was displayed on

TABLE 2 Mean word frequency and length of stimuli (mean and standard deviation).

Word	Directions	Appear form	Frequency (per million)	Word length
Chinese	E-C	Target Chinese translation	3.6 (1.04)	2 (0)
		Written misleading Chinese	3.6 (1.16)	2 (0)
	C-E	Audio presented Chinese	3.5 (0.93)	2 (0)
English	E-C	Audio presented English	3.6 (0.90)	7 (1.61)
	C-E	Target English translation	3.5 (0.90)	7 (1.81)
		Written misleading English	3.2 (0.40)	7 (1.86)

TABLE 3 Descriptive analysis of study abroad (SA) and non-study abroad (NSA) groups' working memory spans and L2 proficiency (mean and standard deviation).

	SA	NSA
Working memory	25.0 (5.90)	24.4 (5.46)
L2 Proficiency	63.1 (7.91)	65.5 (10.73)

the screen. Participants were instructed to rapidly and accurately identify whether the word on the screen corresponded to the correct translation of the auditorily presented word by pressing the Yes or No buttons on the Chronos box. No time limit was imposed on participants' responses.

Two word characteristics—word frequency and word length—were carefully controlled across stimuli and target translation equivalents in both translation directions (L1-L2 and L2-L1) to ensure the generalisability of our findings.

Chinese and English word frequencies were obtained from the SUBTLEX-UK (Van Heuven et al., 2014) and SUBTLEX-CH (Cai and Brysbaert, 2010) word frequency databases, respectively. As demonstrated in Table 2, Chinese and English words used in the task were matched for word frequency. This similarity of frequency and word length indicates that any differences in accuracy or response time during task performance are not attributable to disparities in word frequency between the misleading words and target translation equivalents, but rather are a true reflection of participants' lexical translation performance.

Nonetheless, matching word length across Chinese and English is exceedingly challenging, given their typological distinctions as languages. This study only ensured that the Chinese words employed in both directions comprised a comparable number of characters, while the lengths of the English words utilised in the task were similarly consistent.

Participants' accuracy (%) and response times (ms) were calculated, with only the response times of accurate trials incorporated into the data analysis. Although translation has long been a conventional pedagogical task in L2 and interpreting training, employing a response time-based translation task as a research instrument remains relatively novel (Issa and Shyamala, 2021).

3.2.4. Post-task interview

To delve deeper into participants' subjective experiences and cognitive reflections on lexical translation performance, we conducted a post-task interview. This complementary approach aimed to understand both their self-evaluations of the lexical translation and the underlying reasons for their perceptions.

The interview served two primary purposes:

Firstly, we provided a quantitative self-assessment, allowing participants to rate their performance on a five-point scale (1 = very poor, 5 = excellent).

Secondly, we sought qualitative insights through an open-ended question. This was designed to reveal the underlying factors or thought processes influencing their ratings, providing a deeper understanding of their performance beyond the quantitative assessment.

All interviews were conducted in their native language to ensure comfort and clarity in communication.

3.3. Procedure

Participants consented to the research conducted in a quiet university space using a Lenovo laptop and E-Prime Professional 2.0 software. All verbal instructions were delivered in their native language. Tasks involved word translation recognition and reading span, with response times and accuracy recorded via a Chronos response box. The task order was counterbalanced to mitigate fatigue effects. The session, approximately 30 min, was audio-recorded. It included consent, LexTALE completion, memory span task, bidirectional word translation tasks, and a post-task interview.

4. Results

Table 3 shows the descriptive analysis of participants' working memory spans and L2 proficiency, skewness and kurtosis indicate that the means are normally distributed.

Two groups were comparable in their mean working memory, $t(24) = 0.30$, $p = 0.77$. and their L2 proficiency assessed by LexTALE, $t(24) = 0.85$, $p = 0.41$.

4.1. Overall direction-dependent asymmetry in accuracy

The impact of SAE on the groups' bidirectional word translation was explored via a 2 × (2) mixed factorial ANOVA with the within-subjects factor of the direction of translation (L2-L1 vs. L1-L2) and the between-subjects factor of the group (NSA vs. SA). A significant main effect of direction was observed in accuracy, $F(1, 48) = 35.02$, $p < 0.001$, $\eta_p^2 = 0.422$, as depicted in Table 4 and Figure 2. This indicated that both groups exhibited the same direction-dependent asymmetry: word translation from the L2-L1 direction was more accurate than in the opposite direction.

4.1.1. Comparison of study abroad and non-study abroad groups' accuracy

According to the 2 × (2) mixed factorial ANOVA, a significant main effect of group was observed in the two groups' accuracy, F(1,

TABLE 4 Word recognition accuracy from Chinese to English (L1-L2) and English to Chinese (L2-L1; mean and standard deviation).

	SA	NSA
L1-L2 (% correct)	65% (0.08)	72% (0.08)
L2-L1 (% correct)	70% (0.10)	77% (0.10)

TABLE 5 Word recognition reaction times from Chinese to English (L1-L2) and English to Chinese (L2-L1; mean and standard deviation).

	SA	NSA
L1-L2 (ms)	948.7 (193.95)	1167.9 (193.95)
L2-L1 (ms)	929.9 (214.25)	1032.5 (205.76)

48) = 8.51, $p = 0.005$, $\eta_p^2 = 0.151$, which suggests that the NSA group is more accurate than the SA group in retrieving words in both directions (mean accuracy rate: 74% vs. 68%, respectively).

4.1.2. Interaction between group accuracy and translation direction

According to the 2 × (2) mixed factorial ANOVA, the group × direction interaction was not significant in both groups' accuracy, $F(1, 48) = 0.05$, $p = 0.83$, $\eta_p^2 = 0.001$. This finding suggests that there is no statistically significant difference in the interaction between language direction (L1-L2 or L2-L1) and lexical translation accuracy among the two groups (SA and NSA). In more precise terms, SAE does not result in a discernible variation in the accuracy with which students execute lexical translation tasks across language directions in comparison to their counterparts without such experience.

4.2. Overall direction-dependent asymmetry in response times

The impact of SAE on the groups' bidirectional word translation was explored via a 2 × (2) mixed factorial ANOVA with the within-subjects factor of the direction of translation (L2-L1 vs. L1-L2) and the between-subjects factor of the group (NSA vs. SA). A significant main effect of direction was also observed in response times, $F(1, 48) = 8.96$, $p = 0.004$, $\eta_p^2 = 0.157$, as depicted in Table 5 and Figure 3. This indicated that both groups exhibited the same direction-dependent asymmetry: word translation from the L2-L1 direction was more efficient than in the opposite direction.

4.2.1. Comparison of study abroad and non-study abroad groups' response times

According to the 2 × (2) mixed factorial ANOVA, a significant main effect of group was observed in response times, $F(1, 48) = 9.05$, $p = 0.004$, $\eta_p^2 = 0.159$. This main effect indicates that the SA group demonstrates faster response times than the NSA group in accurately retrieving words in both directions (939.3 ms vs. 1,100.2 ms, respectively). This is evidenced by the shorter response times in the SA group. Given that the SA group exhibited faster response times yet lower accuracy than the NSA group, this phenomenon may be indicative of a speed-accuracy trade-off.

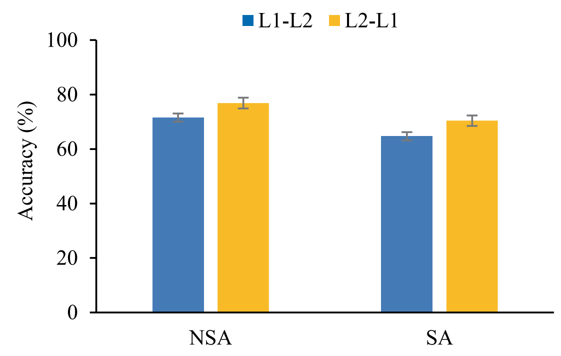


FIGURE 2
Bar graph showing the overall word translation recognition accuracy (in percentage) in Chinese-English (L1-L2) and English-Chinese (L2-L1) directions across both groups regardless of language environment.

4.2.2. Interaction between group response times and translation direction

According to the 2 × (2) mixed factorial ANOVA, a significant group × direction interaction was observed in response times, $F(1, 48) = 5.134$, $p = 0.028$, $\eta_p^2 = 0.097$. To further investigate this interaction, the authors conducted pairwise comparisons using an adjusted alpha level of 0.025. The NSA group exhibited a significant difference in their word translation recognition response times for L1-L2 vs. L2-L1, $t(24) = 4.17$, $p < 0.001$. In contrast, no statistical difference was detected for the SA group, $t(24) = 0.47$, $p = 0.64$, which suggests a more balanced performance across the two language directions among the SA group.

4.2.3. Post-task interview

Table 6 presents the self-rating results of the word translation recognition task. No significant difference was observed between the SA and NSA groups, $t(24) = 0.65$, $p = 0.52$. This indicates that both groups perceived their performance as somewhere around 'average'.

Delving into the qualitative feedback from the open-ended question, a noteworthy pattern emerged. A substantial 17 out of 25 SA participants felt they had rushed through the task. They expressed a post-submission realisation of their errors, indicating a premature commitment to answers. One SA participant reported, 'Although I thought the task would be challenging at first, I later found it to be less difficult than I had anticipated. Nevertheless, I still felt compelled to respond as quickly as possible, maybe due to some external pressure or personal drive'. In contrast, only 2 out of 25 NSA participants reported similar experiences, with most commenting on the difficulty of the task, such as confusing distractors that resembled the correct words.

5. Discussion

5.1. Descriptive analysis of working memory and L2 proficiency

Based on the results presented in Table 3, there appears to be no significant difference in two working memory and L2 proficiency between the SA and NSA groups. This suggests that the two groups

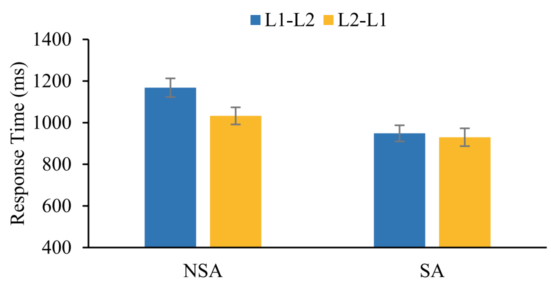


FIGURE 3
Graph depicting the overall word translation recognition response times (in milliseconds) in Chinese-English (L1-L2) and English-Chinese (L2-L1) directions for both groups.

TABLE 6 Participants' self-rating on word translation recognition performance.

	SA	NSA
Self-rating	3.16 (0.62)	3.04 (0.73)

are comparable in terms of these cognitive and linguistic measures, ensuring that any differences observed in other measures are less likely to be attributed to discrepancies in working memory or baseline L2 proficiency.

5.2. Overall direction-dependent asymmetry in accuracy and response times

As illustrated in Table 4 and Figure 2, both groups' word translation from the L2-L1 direction was more accurate and faster than in the opposite direction. This result replicates previous research findings that L1-L2 lexical translation is slower and more error-prone than L2-L1 (e.g., Meuter and Allport, 1999; Costa and Santesteban, 2004; Linck et al., 2008; Meuter, 2009; Peeters et al., 2014; Olson, 2017). The finding also lends support to the Inhibitory Control model (Green, 1998), which argues that in unbalanced bilinguals, the dominant L1 is more active, requiring greater cognitive effort to suppress during language processing. The model states that L1-L2 word translation entails inhibiting the dominating L1 competitors to ensure that the intended L2 words are chosen for output while both languages are activated in a non-selective way (see Green, 1998; Kroll et al., 2008). L1-L2 direction thus demands more mental effort, realised as lower accuracy and longer response times, than suppressing the comparatively weaker L2 in the opposite translation direction.

What is noteworthy is the consistency of this direction-dependent asymmetry across both groups, regardless of their SAE. This consistency echoes findings by Meuter and Allport (1999), which proposed that bilinguals' translational asymmetry is primarily influenced by the proficiency of their languages. Even though SAE contributes to inhibiting L1 dominance and facilitating L2 processing in the current study, it appears that for our group of late unbalanced bilinguals, a three-year SAE was insufficient to counteract L1 dominance. Additionally, all our participants were interpreting students who routinely switch between two languages. This habitual language-switching may have counteracted the attenuation impact on their L1 in the SAE. Consequently, both SA and NSA groups demonstrated the same direction-dependent asymmetry in their word translation recognition tasks. These findings underscore the intricate interplay of language proficiency, immersion, and cognitive control in bilingual contexts.

5.3. Group differences in accuracy and response times

A key differentiator between the two groups in this study is the SAE. The SA group demonstrated faster response times but lower accuracy in lexical translation compared to the NSA group, suggesting a speed-accuracy trade-off. To shed light on these observed differences, we integrated insights from post-task interviews and LEAP-Q results.

5.3.1. Post-task interview insights

Though self-ratings indicated similar self-perceptions of performance across both groups (Table 6), qualitative feedback highlighted differing lexical translation approaches. Notably, a significant portion of the SA group felt they had rushed through the task and later realised their mistakes indicating a potential inclination towards prioritising speed over accuracy. This inclination could be influenced by their experiences in real-life scenarios where rapid responses were essential. This focus on speed was further echoed in remarks from SA participants, underscoring the potential conditioning effect of real-time English interactions.

5.3.2. Language experience and proficiency questionnaire results

The SA group engaged more frequently with English speakers on a weekly basis than the NSA group. Given that regular L2 usage can bolster learners' communicative willingness (Dewaele, 2019), SA participants may exhibit a reduced stringency when selecting translation equivalents, potentially sacrificing precision. Conversely, NSA participants appeared more methodical, prioritising accuracy and delving deeper into the task's challenges.

Our observations align with research by DeKeyser (1991) and Tokowicz et al. (2004), which emphasised SAE's impact on lexical translation accuracy. They noted that individuals with extended SAE (exceeding 1 year) were more likely to guess unknown word meanings, despite error risks, contrasting with those with limited or no SAE.

The potential role of cognitive control also merits attention. Previous research highlights the intricate relationship between language use experience and cognitive control abilities (Green and Abutalebi, 2013). Xie and Dong (2017) found that individuals with public-speaking training experience exhibited quicker response times than both monolinguals and a control bilingual group. Similarly, continuous L2 engagement, characterised by intensive semantic, attentional demands, and unwanted behaviour or word suppression, could potentially sharpen cognitive control, as found in studies on public speakers. Our SA group's frequent English interactions could be seen as a catalyst for their enhanced cognitive control, which might manifest in swifter response times.

It is crucial to differentiate between working memory and cognitive control. While working memory manages information retention and manipulation, essential for tasks like language processing (Baddeley, 2017), cognitive control oversees and coordinates cognitive processes, especially amidst distractions. This includes functions such as attentional focus, inhibition, and task switching, vital in bilingual contexts.

In this study's context, both SA and NSA groups having similar working memory capacities suggests equivalent foundational cognitive capabilities. Yet, differences in their SAE could have uniquely moulded their cognitive control faculties. Regular immersion in L2, as experienced by the SA group, intensifies cognitive control demands due to continuous language switching, L1 inhibition, and tackling L2 challenges. Such immersion could have finetuned the SA group's cognitive control, even if their working memory remained consistent with the NSA group.

This variance may explain the observed differences in cognitive control between the two groups, despite comparable working memory capacities. It is not a claim of cognitive superiority but indicates that linguistic experiences might influence cognition differently. Future studies should delve deeper into this by evaluating working memory and cognitive control in comparable bilingual cohorts.

5.4. Interaction between group and translation direction

The interaction effects in the context of accuracy and response times yielded contrasting insights. While there was no interaction effect for accuracy (indicating that both groups showed a similar pattern of direction-dependent asymmetry in accuracy), there was a significant interaction for response times. The NSA group displayed a notable difference in their response times for L1-L2 vs. L2-L1, whereas the SA group did not.

These findings imply that having an SAE might lead to a more consistent and balanced performance in word translation across both language directions. In contrast, those without such an experience may face more variability in their translation speeds depending on the direction.

Our findings replicate that of Schwartz and Kroll (2006) and Chmiel (2016) that the SA group was more balanced and consistent in terms of lexical processing during word translation than the NSA group. Previous studies have suggested that factors such as participants' L2 proficiency (Meuter and Allport, 1999; Costa and Santesteban, 2004; Costa, 2005), language-switching habits (Christoffels et al., 2007), working memory (Sunderman and Kroll, 2009), and learning contexts (Kroll et al., 1998; Kroll and Sunderman, 2003; Linck et al., 2008) impact on the language inhibitory process, and influence the efforts involved in lexical processing. In the present study, participants in the SA and NSA groups were all interpreting students, and therefore, all were engaged in bilingual processing on a daily basis. Moreover, they were also comparable in their L2 proficiency, word knowledge and working memory resource availability. Therefore, the smaller degree of asymmetry observed in the SA group may be attributed to their habitual toggling between two languages during SAE.

For the SA group, immersion in an L2-rich environment granted them a distinct advantage by allowing them to suppress interference from their dominant L1 more effectively (e.g., Linck et al., 2009; Baus et al., 2013). This reduced effort in inhibiting the L1 during L2 processing manifested even though their L1 remained dominant. Consequently, the SA group exhibited more consistent and less asymmetric performance in bilingual processing compared to their NSA counterparts.

It is well-established that prolonged exposure to bilingual environments sharpens a bilingual's ability to switch languages (e.g., Bialystok and Barac, 2012; Nicolay and Poncelet, 2015). Moreover, frequent daily engagements with L2, as observed in study-abroad bilinguals, not only diminish the influence of L1 but also enhance the ease of transitioning between both languages (Bonfieni et al., 2019). Echoing Xie and Dong (2021), such bilinguals predominantly engage in English, particularly when interacting with peers, making linguistic toggling commonplace. This frequent language transition, characteristic of the SA group, underlines their improved mental set shifting, leading to a more balanced bilingual lexical translation compared to their NSA counterparts.

In contrast, for NSA participants in our study, their daily linguistic environment was predominantly aligned with their native L1. This might have heightened the challenge of suppressing the ever-present L1 during L2 processing, thus skewing their bilingual lexical translation. Moreover, their interactions in English were notably fewer than those of the SA group, as indicated by their LEAP-Q results. This suggests that they predominantly communicated in Chinese, especially with fellow Chinese students, limiting their opportunities for smooth transitions between languages. The absence of a consistent L2-rich environment might deprive the NSA group of the routine that aids mental set shifting, especially the transition between languages. Consequently, this might lead to a pronounced degree of asymmetry in the NSA group's performance.

Furthermore, it is pertinent to underscore the unique linguistic profile of interpreting students. As trainees navigating two languages on a daily basis, they embody a distinct category of bilinguals. Language switching is integral to their training, sharpening their ability to transition between languages rapidly. Even amidst this frequent toggling, the impact of SAE on lexical processing emerged prominently in our findings. Such a pronounced effect, in spite of their rigorous linguistic exercises as interpreting students, further underscores the profound influence of SAE. It suggests that while routine interpreting practices equip students with certain bilingual proficiencies and cognitive advantages, immersion in an authentic language environment through SAE offers unparalleled benefits.

To wrap up our discussion, our findings indicate that concerning direction-dependent asymmetry, both SA and NSA groups were more adept in translating words from L2 to L1 than vice versa. Although we postulated that the SA group showcased more efficient bidirectional word translation, the results were multifaceted. While the SA participants were faster across translation directions, the NSA group exhibited greater accuracy. This divergence hints at a speed-accuracy trade-off: SA participants, due to their extensive L2 usage and frequent language toggling, might prioritise speed and a propensity to communicate, even without the precise word.

Moreover, our research underlines that SAE aids in achieving a less asymmetric performance in word translations across languages. Such observations emphasise the instrumental role of SAE in alleviating language competition, diminishing the cognitive strain tied to bilingual lexical processing, and fine-tuning the cognitive mechanism managing bilingualism.

This study is not without its limitations, notably its cross-sectional design. Adopting a longitudinal methodology, tracking the same group of students during their SAE might offer richer insights into the progression of their lexical translation performance.

Additionally, we assessed only the overarching working memory resources, bypassing specific facets of cognitive control. Prospective research should delve into this, appraising both working memory and cognitive control in comparable bilingual groups.

Lastly, the SAE is a composite of linguistic immersion and the intricacies of residing in a foreign country. While our participants' increased willingness to communicate might arise from intensified linguistic exposure and interaction in a study-abroad context, it could also be shaped by non-linguistic elements like the process of cultural adaptation. As Xie and Dong (2021) have highlighted, the confluence of these components—linguistic immersion, cultural adjustments, and other unique challenges faced abroad—might collectively influence bilingual performance and communicative behaviour. This complex interplay undeniably warrants further exploration.

6. Conclusion

This study elucidates the impact of SAE on bidirectional lexical translation among Chinese (L1) English interpreting students. While previous research has touched upon the effects of SAE on bilingual translation, our findings augment this body of knowledge by highlighting the performance differences between the SA and NSA groups. Notably, the SA group showcased superior consistency in their translations and displayed heightened communicative willingness.

Recognising the crucial role of lexical processing in higher-order language processing, including interpreting, there's an evident need for pedagogical adjustments. We advocate for universities to bolster communicative activities both within and beyond the curriculum, thereby immersing students more deeply in their L2. Such active engagement can potentiate the activation of their L2, mitigating the cognitive burdens of L2 processing and language switching. This approach bears significant relevance for interpreting learners, aiding them in honing critical skills for their academic and future professional endeavours.

While numerous studies have delved into bilingualism across varied language-learning contexts, there remains a paucity of research focusing on the SAE's impact on the Chinese-English language pairs, especially among interpreting students. As such, our results not only bridge this gap but also furnish actionable insights for L2 educators and interpreting trainers. The findings are especially pertinent for interpreting students without the advantage of SAE, offering them strategies to compensate for their limited interactions in L2 environments.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving humans were approved and conducted by Western Sydney University Human Research Ethics Committee's guidelines (approval number H12405). The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

RW: Conceptualization, Data curation, Formal Analysis, Investigation, Writing – original draft, Writing – review & editing. JH: Supervision, Writing – review & editing. BB: Supervision, Writing – review & editing. MA: Data curation, Formal Analysis, Software, Supervision, Writing – review & editing.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

Conflict of interest

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

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Antoniou, M., Wong, P. C., and Wang, S. (2015). The effect of intensified language exposure on accommodating talker variability. *J. Speech Lang. Hear. Res.* 58, 722–727. doi: 10.1044/2015_JSLHR-S-14-0259

Ardila, A. (2003). Language representation and working memory with bilinguals. *J. Commun. Disord.* 36, 233–240. doi: 10.1016/S0021-9924(03)00022-4

Babcock, L., and Vallesi, A. (2017). Are simultaneous interpreters expert bilinguals, unique bilinguals, or both? *Biling. Lang. Congn.* 20, 403–417. doi: 10.1017/S1366728915000735

Baddeley, A. (1996). Exploring the central executive. *Quart. J. Experiment. Psychol. Sec. A* 49, 5–28. doi: 10.1080/713755608

Baddeley, A. (2017). Modularity, working memory and language acquisition. *Second. Lang. Res.* 33, 299–311. doi: 10.1177/0267658317709852

Barquin, E. L. (2012). Writing development in a study abroad context [Doctoral dissertation]. Universitat Pompeu Fabra.

Bartolotti, J., Marian, V., Schroeder, S. R., and Shook, A. (2011). Bilingualism and inhibitory control influence statistical learning of novel word forms. *Front. Psychol.* 2:324. doi: 10.3389/fpsyg.2011.00324

Baus, C., Costa, A., and Carreiras, M. (2013). On the effects of second language immersion on first language production. *Acta Psychol.* 142, 402–409. doi: 10.1016/j.actpsy.2013.01.010

Berghoff, R., McLoughlin, J., and Bylund, E. (2021). L1 activation during L2 processing is modulated by both age of acquisition and proficiency. *J. Neurolinguistics* 58:100979. doi: 10.1016/j.jneuroling.2020.100979

Bhandari, R. (2017). A world on the move: trends in global student mobility (higher education policy for minorities in the United States) [press release]. IIE Center for Academic Mobility Research and Impact. Available at: http://hdl.handle.net/10919/83142

Bialystok, E., and Barac, R. (2012). Emerging bilingualism: dissociating advantages for metalinguistic awareness and executive control. *Cognition* 122, 67–73. doi: 10.1016/j.cognition.2011.08.003

Bialystok, E., Craik, F. I., Green, D. W., and Gollan, T. H. (2009). Bilingual minds. *Psychol. Sci. Public Interest* 10, 89–129. doi: 10.1177/1529100610387084

Bonfieni, M., Branigan, H. P., Pickering, M. J., and Sorace, A. (2019). Language experience modulates bilingual language control: the effect of proficiency, age of acquisition, and exposure on language switching. *Acta Psychol.* 193, 160–170. doi: 10.1016/j.actpsy.2018.11.004

Cai, Q., and Brysbaert, M. (2010). SUBTLEX-CH: Chinese word and character frequencies based on film subtitles (research article). *PLoS One* 5:e10729. doi: 10.1371/journal.pone.0010729

Chmiel, A. (2016). Directionality and context effects in word translation tasks performed by conference interpreters. *Poznan Stud. Contemporary Linguist.* 52, 269–295. doi: 10.1515/psicl-2016-0010

Christoffels, I., De Groot, A. M. B., and Waldorp, L. J. (2003). Basic skills in a complex task: a graphical model relating memory and lexical retrieval to simultaneous interpreting. *Biling. Lang. Congn.* 6, 201–211. doi: 10.1017/s1366728903001135

Christoffels, I., Firk, C., and Schiller, N. (2007). Bilingual language control: an event-related brain potential study. *Brain Res.* 1147, 192–208. doi: 10.1016/j.brainres.2007.01.137

Costa, A. (2005). Lexical access in bilingual production. In GrootA. M. B. De and J. Kroll (Eds.), *Handbook of bilingualism: psycholinguistic approaches* (pp. 389–407). Oxford: Oxford University Press.

Costa, A., Hernández, M., Costa-Faidella, J., and Sebastián-Gallés, N. (2009). On the bilingual advantage in conflict processing: now you see it, now you don't. *Cognition* 113, 135–149. doi: 10.1016/j.cognition.2009.08.001

Costa, A., and Santesteban, M. (2004). Lexical access in bilingual speech production: evidence from language switching in highly proficient bilinguals and L2 learners. *J. Memory* 50, 491–511. doi: 10.1016/j.jml.2004.02.002

Daneman, M., and Carpenter, P. A. (1980). Individual differences in working memory and reading. *J. Verbal Learn. Verbal Behav.* 19, 450–466. doi: 10.1016/S0022-5371(80)90312-6

Daneman, M., and Merikle, P. M. (1996). Working memory and language comprehension: a meta-analysis. *Psychon. Bull. Rev.* 3, 422–433. doi: 10.3758/BF03214546

De Bruin, A., Roelofs, A., Dijkstra, T., and FitzPatrick, I. (2014). Domain-general inhibition areas of the brain are involved in language switching: FMRI evidence from trilingual speakers. *NeuroImage* 90, 348–359. doi: 10.1016/j.neuroimage.2013.12.049

De Groot, A. M., and Poot, R. (1997). Word translation at three levels of proficiency in a second language: the ubiquitous involvement of conceptual memory. *Lang. Learn.* 47, 215–264. doi: 10.1111/0023-8333.71997007

DeKeyser, R. (1991). The semester overseas: what difference does it make. *ADFL Bulletin* 22, 42–48.

DeLuca, V., Segaert, K., Mazaheri, A., and Krott, A. (2020). Understanding bilingual brain function and structure changes? U bet! A unified bilingual experience trajectory model. *J. Neurolinguistics* 56:100930. doi: 10.1016/j.jneuroling.2020.100930

Dewaele, J.-M. (2019). The effect of classroom emotions, attitudes toward English, and teacher behavior on willingness to communicate among English foreign language learners. *J. Lang. Soc. Psychol.* 38, 523–535. doi: 10.1177/0261927X19864996

Fitzpatrick, T. (2012). Tracking the changes: vocabulary acquisition in the study abroad context. *Lang. Learn. J.* 40, 81–98. doi: 10.1080/09571736.2012.658227

Gile, D. (2009). *Basic concepts and models for interpreter and translator training* (*Rev. ed.*). Netherlands: John Benjamins.

Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Biling. Lang. Congn.* 1, 67–81. doi: 10.1017/S1366728998000133

Green, D. W., and Abutalebi, J. (2013). Language control in bilinguals: the adaptive control hypothesis. *J. Cogn. Psychol.* 25, 515–530. doi: 10.1080/20445911.2013.796377

Grey, S., Cox, J. G., Serafini, E. J., and Sanz, C. (2015). The role of individual differences in the study abroad context: cognitive capacity and language development during short-term intensive language exposure. *Mod. Lang. J.* 99, 137–157. doi: 10.1111/modl.12190

Grosjean, F. (1997). The bilingual individual. *Interpreting* 2, 163–187. doi: 10.1075/intp.2.1-2.07gro

Hamano-Bunce, D., Murray, R., and Campbell, B. (2019). The effects of a short study abroad Programme on Japanese learners' L2 listening. *Asian EFL J. Quart.* 23, 106–129.

Han, S., and Ma, Y. (2014). Cultural differences in human brain activity: a quantitative meta-analysis. *NeuroImage* 99, 293–300. doi: 10.1016/j.neuroimage.2014.05.062

Issa, S. H., Awadh, F. H., and Ahmed, H. R. (2022). The role of proficiency level in the speed of lexical activation. *Cogent Arts Human.* 9:1999613. doi: 10.1080/23311983.2021.1999613

Issa, B. I., Faretta Stutenberg, M., and Bowden, H. W. (2020). Grammatical and lexical development during short-term study abroad: exploring L2 contact and initial proficiency. *Mod. Lang. J.* 104, 860–879. doi: 10.1111/modl.12677

Issa, S. H., and Shyamala, K. C. (2021). Investigating the speed of lexical activation in the bilinguals' L1 and L2 through forward and backward translation task. *Int. J. English Lang. Stud.* 3, 26–34. doi: 10.32996/ijels.2021.3.3.4

Jarvis, S., and Pavlenko, A. (2008). *Crosslinguistic influence in language and cognition*. England: Routledge.

Kaushanskaya, M., Yoo, J., and Marian, V. (2011). The effect of second-language experience on native-language processing. *Vigo Int. J. App. Linguis.* 8, 54–77.

Keuleers, E., Stevens, M., Mandera, P., and Brysbaert, M. (2015). Word knowledge in the crowd: measuring vocabulary size and word prevalence in a massive online experiment. *Q. J. Exp. Psychol.* 68, 1665–1692. doi: 10.1080/17470218.2015.1022560

Kleinman, D., Morgan, A. M., Ostrand, R., and Wittenberg, E. (2022). Lasting effects of the COVID-19 pandemic on language processing. *PLoS One* 17:e0269242. doi: 10.1371/journal.pone.0269242

Kroll, J., Bobb, S. C., Misra, M., and Guo, T. (2008). Language selection in bilingual speech: evidence for inhibitory processes. *Acta Psychol.* 128, 416–430. doi: 10.1016/j.actpsy.2008.02.001

Kroll, J., Michael, E. B., and Sankaranarayanan, A. (1998). "A model of bilingual representation and its implications for second language acquisition" in *Foreign language learning: Psycholinguistic studies on training and retention*. eds. A. Healy and L. E. Bourne (United States: Lawrence Erlbaum Associates), 365–395.

Kroll, J., and Stewart, E. (1994). Category interference in translation and picture naming: evidence for asymmetric connections between bilingual memory representations. *J. Mem. Lang.* 33, 149–174. doi: 10.1006/jmla.1994.1008

Kroll, J., and Sunderman, G. (2003). "Cognitive processes in second language learners and bilinguals—the development of lexical and conceptual representation" in *The handbook of second language acquisition*. eds. C. J. Doughty and M. Long (Hoboken: Blackwell Publishing), 104–129.

Lara, A. R. (2014). Complexity, accuracy and fluency development through study abroad programmes varying in duration [Doctoral dissertation]. Universitat Pompeu Fabra.

Lemhöfer, K., and Broersma, M. (2012). Introducing LexTALE: a quick and valid lexical test for advanced learners of English. *Behav. Res. Ther.* 44, 325–343. doi: 10.3758/s13428-011-0146-0

Linck, J. A., Hoshino, N., and Kroll, J. (2008). Cross-language lexical processes and inhibitory control. *Mental Lexicon* 3, 349–374. doi: 10.1075/ml.3.3.06lin

Linck, J. A., Kroll, J., and Sunderman, G. (2009). Losing access to the native language while immersed in a second language: evidence for the role of inhibition in second-language learning. *Psychol. Sci.* 20, 1507–1515. doi: 10.1111/j.1467-9280.2009.02480.x

Llanes, À., and Muñoz, C. (2009). A short stay abroad: does it make a difference? *System* 37, 353–365. doi: 10.1016/j.system.2009.03.001

Mead, P. (2002). "Exploring hesitation in consecutive interpreting: an empirical study" in *Interpreting in the 21st century*. eds. G. Garzone and M. Viezzi (Netherlands: John Benjamins), 73–82.

Meuter, R. F. (2009). Language selection and performance optimisation in multilinguals. In L. Isurin, D. Winford and BotK. De (Eds.), *Multidisciplinary approaches to code switching* (pp. 27–51). Netherlands: John Benjamins.

Meuter, R. F., and Allport, A. (1999). Bilingual language switching in naming: asymmetrical costs of language selection. *J. Mem. Lang.* 40, 25–40. doi: 10.1006/jmla.1998.2602

Miyake, A., and Friedman, N. P. (1998). "Individual differences in second language proficiency: working memory as language aptitude" in *Foreign language learning: Psycholinguistic studies on training and retention*. eds. A. Healy and L. E. Bourne (United States: Lawrence Erlbaum Associates), 339–364.

Muñoz, C. (2008). Symmetries and asymmetries of age effects in naturalistic and instructed L2 learning. *Appl. Linguis.* 29, 578–596. doi: 10.1093/applin/amm056

Nicolay, A.-C., and Poncelet, M. (2015). Cognitive benefits in children enrolled in an early bilingual immersion school: a follow up study. *Biling. Lang. Congn.* 18, 789–795. doi: 10.1017/S1366728914000868

O'Brien, I., Segalowitz, N., Collentine, J., and Freed, B. (2006). Phonological memory and lexical, narrative, and grammatical skills in second language oral production by adult learners. *Appl. Psycholinguist.* 27, 377–402. doi: 10.1017/S0142716406060322

Odlin, T. (2003). "Cross linguistic influence" in *The handbook of second language acquisition*. eds. C. J. Doughty and M. Long (Hoboken: Blackwell Publishing), 436–486.

Odlin, T. (2012). "Crosslinguistic influence in second language acquisition" in *The encyclopedia of applied linguistics*. ed. C. A. Chapelle (Hoboken: Blackwell Publishing), wbeal0292.

Olson, D. (2017). Bilingual language switching costs in auditory comprehension. *Lang. Cogn. Neurosci.* 32, 494–513. doi: 10.1080/23273798.2016.1250927

Pavlacic, P. (2018). Raising the profile in China of Australia's excellence in the delivery of English language training [press release]. English Australia. Available at: https://internationaleducation.gov.au/International-network/Australia/InternationalStrategy/EGIProjects/Documents/Raising%20the%20Profile%20in%20China%20in%20English%20Language%20Training%20FINAL.pdf

Peeters, D., Runnqvist, E., Bertrand, D., and Grainger, J. (2014). Asymmetrical switch costs in bilingual language production induced by reading words. *J. Exp. Psychol. Learn. Mem. Cogn.* 40:284. doi: 10.1037/a0034060

Pérez-Vidal, C., Juan-Garau, M., Mora, J. C., and Valls-Ferrer, M. (2012). "Oral and written development in formal instruction and study abroad: differential effects of learning context" in *Intensive exposure experiences in second language learning*. ed. C. Muñoz, vol. *65* (Bristol: Multilingual Matters), 213–233.

Prior, A., and Gollan, T. H. (2011). Good language-switchers are good task-switchers: evidence from Spanish–English and mandarin–English bilinguals. *J. Int. Neuropsychol. Soc.* 17, 682–691. doi: 10.1017/S1355617711000580

Ransdell, S., Barbier, M.-L., and Niit, T. (2006). Metacognitions about language skill and working memory among monolingual and bilingual college students: when does multilingualism matter? *Int. J. Biling. Educ. Biling.* 9, 728–741. doi: 10.2167/beb390.0

Russell, D., and Takeda, K. (2015). "Consecutive interpreting" in *The Routledge handbook of interpreting*. eds. H. Mikkelson and R. Jourdenais (England: Routledge), 96–111.

Santilli, M., Vilas, M. G., Mikulan, E., Caro, M. M., Muñoz, E., Sedeño, L., et al. (2019). Bilingual memory, to the extreme: lexical processing in simultaneous interpreters. *Biling. Lang. Congn.* 22, 331–348. doi: 10.1017/S1366728918000378

Schwartz, A. I., and Kroll, J. (2006). Bilingual lexical activation in sentence context. *J. Memory Lang. Cogn. Process.* 55, 197–212. doi: 10.1016/j.jml.2006.03.004

Segalowitz, N., and Freed, B. (2004). Context, contact, and cognition in oral fluency acquisition: learning Spanish in at home and study abroad contexts. *Stud. Second. Lang. Acquis.* 26, 173–199. doi: 10.1017/S0272263104262027

Signorelli, T., Haarmann, H. J., and Obler, L. K. (2012). Working memory in simultaneous interpreters: effects of task and age. *Int. J. Biling.* 16, 198–212. doi: 10.1177/1367006911403200

Sunderman, G., and Kroll, J. (2009). When study-abroad experience fails to deliver: the internal resources threshold effect. *Appl. Psycholinguist.* 30, 79–99. doi: 10.1017/S0142716408090048

Szego, J. (2020). Australia's Chinese lesson—the nation's urgent need to engage with international students from China (press release NARRATIVE #4). Scanlon Foundation Research Institute. Available at: https://scanloninstitute.org.au/publication/australias-chinese-lesson-nations-urgent-need-engage-international-students-china

Tokowicz, N., Michael, E. B., and Kroll, J. (2004). The roles of study-abroad experience and working-memory capacity in the types of errors made during translation. *Biling. Lang. Congn.* 7, 255–272. doi: 10.1017/S1366728904001634

Tu, L., Wang, J., Abutalebi, J., Jiang, B., Pan, X., Li, M., et al. (2015). Language exposure induced neuroplasticity in the bilingual brain: a follow-up fMRI study. *Cortex* 64, 8–19. doi: 10.1016/j.cortex.2014.09.019

Van Heuven, W., Mandera, P., Keuleers, E., and Brysbaert, M. (2014). SUBTLEX-UK: a new and improved word frequency database for British English. *Q. J. Exp. Psychol.* 67, 1176–1190. doi: 10.1080/17470218.2013.850521

Xie, Z., and Dong, Y. (2017). Contributions of bilingualism and public speaking training to cognitive control differences among young adults. *Biling. Lang. Congn.* 20, 55–68. doi: 10.1017/S1366728915000474

Xie, Z., and Dong, Y. (2021). Influence of the study abroad bilingual experience on cognitive control among young adults. *Int. J. Biling.* 25, 1417–1428. doi: 10.1177/13670069211023126

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 12 October 2023
DOI 10.3389/fpsyg.2023.1270743

Check for updates

OPEN ACCESS

EDITED BY
Monika M. Polczynska,
University of California, Los Angeles,
United States

REVIEWED BY
Kaile Zhang,
University of Macau, China
Vincent Van Heuven,
Leiden University, Netherlands

*CORRESPONDENCE
Hanna Kędzierska
✉ hanked@amu.edu.pl

RECEIVED 01 August 2023
ACCEPTED 26 September 2023
PUBLISHED 12 October 2023

CITATION
Kędzierska H, Rataj K, Balas A, Cal Z, Castle C and Wrembel M (2023) Vowel perception in multilingual speakers: ERP evidence from Polish, English and Norwegian. *Front. Psychol.* 14:1270743.
doi: 10.3389/fpsyg.2023.1270743

Vowel perception in multilingual speakers: ERP evidence from Polish, English and Norwegian

Hanna Kędzierska[1,2]*, Karolina Rataj[3], Anna Balas[1], Zuzanna Cal[1], Chloe Castle[4] and Magdalena Wrembel[1]

[1]Department of Contemporary English Language, Faculty of English, Adam Mickiewicz University, Poznań, Poland, [2]Department of English and Comparative Linguistics, Institute of English Studies, University of Wrocław, Wrocław, Poland, [3]Neuroscience of Language Laboratory, Department of Psycholinguistic Studies, Faculty of English, Adam Mickiewicz University, Poznań, Poland, [4]Department of Language and Culture, UiT The Arctic University of Norway, Tromsø, Norway

Introduction: Research on Mismatch Negativity (MMN) in monolingual and bilingual speakers has shown significant differences in L1 versus L2 phonemic perception. In this study, we examined whether the MMN response is sensitive to the differences between L1, L2 and L3/L*n*.

Methods: We compared bioelectrical brain activity in response to changes in pairs of vowels produced in three different languages. Specifically, multilingual participants listened to selected vowel contrasts in their L1 Polish, L2 English and L3/L*n* Norwegian presented within the passive-oddball paradigm.

Results: Results revealed that the MMN was modulated by language: we observed significant differences between L2 English and L3/L*n* Norwegian as well as between L1 Polish and L3/L*n* Norwegian. For L3/L*n* Norwegian, the MMN response had a lower amplitude when compared with L2 English and L1 Polish.

Discussion: Such findings suggest that foreign language status (i.e., L2 vs. L3/L*n*) modulates early auditory processing.

KEYWORDS

multilingualism, speech perception, third language (L3/L*n*), event-related potentials (ERP), mismatch negativity (MMN)

1. Introduction

Non-native phonemic perception is considered a vital component of successful language learning and has become a focus of scientific research. Due to global migration processes and the introduction of at least one foreign language at the early stages of education, multilingualism has become a norm rather than an exception in most European countries. Still, many issues related to the interaction of more than two languages in a single speaker are yet to be investigated, and, among them, those related to the phonemic perception mechanisms (Cabrelli and Wrembel, 2016). The problem, aside from associated theoretical implications, is particularly relevant from the point of view of language learners, who often aim at target-like non-native phoneme production. This intention is very strongly intertwined with their perception of foreign phonemes relative to native phonemes. Previous research has found significant neural differences in native as opposed to non-native phonemic perception, suggesting reduced phonemic discrimination mechanisms in the second language (L2) when compared with the first language (L1) (e.g., Jakoby et al., 2011; Song and Iverson, 2018; Liang and Chen, 2022). However, the listener's auditory discrimination abilities in L3/L*n* remain largely understudied. While proficiency in L2 is generally considered an advantage in acquiring L3/L*n* phonologies, an ongoing scientific debate on multilingualism tends to highlight the complexity of multiple languages interacting in the same speaker (e.g., Wrembel, 2015; Slabakova, 2017; Westergaard et al., 2017; Wrembel and Cabrelli Amaro, 2018;

Dziubalska-Kołaczyk and Wrembel, 2022). Investigating the neural pattern associated with trilingual (as opposed to bilingual) listeners could then greatly contribute to this debate. In the current study we investigated the mismatch negativity (MMN) event-related brain potentials (ERP) response to phonemic differences in participants whose L1 was Polish, L2 was English, and L3/L*n* – Norwegian. Given the previously observed discrepancies in phonemic discrimination in low versus high proficiency language learners (Liang and Chen, 2022), we also accounted for participants' language proficiency and dominance.

The seminal study of Näätänen et al. (1997) revealed that listeners' sensitivity to native phonemes can be indexed by the MMN component, thus beginning a series of studies focused on neural responses to phonemic stimuli. Näätänen et al. (1997) presented two groups of monolingual participants – Estonians and Finns – with vowel phonemes existent in both investigated languages (i.e., /e/ and /ö/) and the vowel /õ/ which has phonemic status in Estonian, but not in Finnish. The paradigm used in this and numerous other studies investigating phonological representations in the brain was the passive-oddball paradigm, where a sequence of frequently occurring *standard* stimuli is interrupted by the occasional appearance of a *deviant* stimulus. The use of the paradigm is frequently combined with the event-related brain potentials (ERP) technique, whose main advantage is its exceptionally high temporal resolution and hence high suitability for studying rapidly occurring cognitive processes, such as those related to language comprehension (Luck, 2005; Kaan, 2007; Cohen, 2014). Ideally, these processes may be further reflected in specific ERP components elicited as a reaction to the experimental manipulation and usually described on the basis of polarity, time of occurrence and scalp distribution.

In oddball tasks, the occurrence of the deviant is associated with the MMN response, i.e., a negative-going wave deflection with frontocentral distribution peaking around 150–250 milliseconds from the onset of the deviant (Kaan, 2007; Näätänen et al., 2007). The MMN, with its generators located in the auditory cortex (Alho et al., 1995), is believed to index auditory discrimination at the pre-attentional level. Thus, its elicitation does not require participants' attention, which may be turned to other types of tasks, such as reading or watching a movie. The MMN is sometimes followed by the P300 component, i.e., a positive deflection observed at around 300 ms after change onset (Polich, 2012). P300 can be further divided into P3a and P3b sub-components, associated, respectively, with attentional switching and memory storage, which differ in terms of latency (with P3a occurring earlier) and distribution (with P3a being more anterior) (Roehm et al., 2007). Another component which has been demonstrated to follow the MMN is late discriminative negativity (LDN), i.e., a negativity observed over frontocentral sites at around 350–600 ms after change onset and typically associated with pre-attentive cognitive evaluation of the stimulus (Ceponiene et al., 1998; Jakoby et al., 2011; Liang and Chen, 2022). In Näätänen et al. (1997), the Estonian group showed an enhanced MMN response when compared with Finnish listeners if the deviant stimulus was /õ/, which has phonemic status only in Estonian. The finding suggested increased neural response to native phonemes and has consequently encouraged debate concerning phonemic discrimination in bilingual speakers.

Notably, further studies investigating phonological sensitivity in bilingual listeners have delivered divergent results, thus implying the importance of listener-oriented factors in the processing of non-native phonemic contrasts. Winkler et al. (1999) observed a similar MMN response to Finnish vowel contrasts in native speakers of Finnish and a group of Hungarians who were late learners of Finnish but who acquired the language to an advanced level in a naturalistic setting. However, in a similar study, Peltola et al. (2003) found a significant difference between native speakers of English and advanced students of English (native speakers of Finnish) who learnt English in a classroom setting. English vowel contrasts evoked lower MMN amplitudes in the latter group. This result seems to be further corroborated by Wottawa et al. (2022), who also found a diminished MMN response in proficient German learners who acquired German at school. The apparent discrepancy in the previously obtained results seems to indicate the importance of the learning context as a vital component of non-native phonemic perception. This hypothesis is additionally supported by the findings of Peltola et al. (2012) who observed that the MMN amplitude in dominant bilinguals depended on the language context of the experiment (i.e., the language used by the experimenter). In the current study, we focused on two foreign languages acquired in two different learning settings: most of our participants started learning English from age 10 onwards in the classroom setting and then migrated to Norway in adulthood, hence learning Norwegian at a later stage in life and in a much more naturalistic way.

Importantly, apart from the context of acquisition and the experimental setting *per se*, another factor which has been demonstrated to affect the pre-attentional phoneme discrimination in L2 is the level of proficiency. Liang and Chen (2022) found different neural responses in adult Mandarin learners of English with high and low proficiency levels. When processing non-native phonemic contrasts, bilinguals with high L2 proficiency showed the MMN response followed by late discriminative negativity (LDN). In contrast, participants with lower L2 proficiency showed the P3b component followed by the late positive component (LPC), i.e., a positivity observed in the parietal region between 250 and 600 ms (Liang and Chen, 2022). This result points to lower proficiency L2 learners' reliance on memory resources in non-native phoneme discrimination. Furthermore, the study of Díaz et al. (2016) demonstrated that MMN was attenuated in poor L2 perceivers, i.e., a group of participants whose vowel contrasts perception was assessed as low in independent behavioral tasks (i.e., a categorization task, a word identification task and a lexical decision task; Díaz et al., 2016: 959). This finding suggests that individual speech-specific capabilities may be a source of variability in L2 phonemic learning.

The main objective of the current research is to shed more light on the perception of non-native phoneme contrasts, and more specifically, to determine whether such contrasts will be equally easy to detect in L2 and L3/L*n*. Testing trilingual listeners should expand the scope of research on both non-native phoneme perception and on multilingualism in general. This way we wished to we go beyond the bilingualism bias which in our opinion does not adequately reflect the current linguistic landscape. What is more, by testing trilingual participants who acquired their non-native languages through distinct modalities and exhibited diverse proficiencies, we wished to disentangle the divergent results of some of the previous studies on non-native phoneme perception in bilinguals.

Specifically, we investigated the perception of L1 Polish /ɨ/−/ɛ/, L2 English /ɪ/−/ʊ/ and L3 Norwegian /i/−/ʏ/ vowel pairs using the ERP technique in a passive oddball paradigm. It should be noted that there are significant differences between the sound systems of the three investigated languages, involving, among other phenomena, the vowel inventory density. While Polish has a fairly scarce vowel repertoire, with only six monophthongal vowels (Jassem, 2003), the vocalic inventories of English and Norwegian are richer with 12 and 18 monophthongal vowels, respectively, (Kristoffersen, 2000; Hawkins and Midgley, 2005; Bjelaković, 2017). The languages differ with respect to combination of lip-rounding with backness. All of them have front unrounded vowels and back rounded vowels, English and Norwegian have high central rounded vowels, whereas only Norwegian has front rounded vowels, which seem to be more marked (i.e., dispreferred among world languages; Maddieson, 2013). In the case of participants in the current study, the order of acquisition would then presume a gradual enlargement of the learners' phonemic (and, specifically, vocalic) repertoire. The above-mentioned phonological differences between the three investigated languages motivated our decision to present vowel contrasts from each language independently, in separate experimental blocks (following previous researchers, e.g., Díaz et al., 2016; Liang and Chen, 2022).

For the sake of comparability of the influence of language status on the processing of native and non-native vowels, an ideal configuration of stimuli would involve the same standard stimulus in all the three languages, and deviants that would be equally distant in terms of all the features from the standard in all the three languages and at the same time these would need to be three different vowels. Such configurations are unattested in real languages; if phones are equidistant and differ with respect to the same features, they are the same sound. If we wanted to compare different vowels in the three languages, we needed to make compromises regarding the degree in which they differed.

Consequently, the choice of standard stimuli was motivated by the high degree of cross-linguistic similarity between the three standard sounds, i.e., the Polish /ɨ/, the English /ɪ/ and the Norwegian /i/ sound. The choice of deviants, on the other hand, was motivated by the systematic differences between the three investigated languages, which were briefly mentioned above. And thus, the Polish /ɨ/—/ɛ/ contrast is mainly manifested in height and is also existent in the other investigated languages. The English /ɪ/—/ʊ/ contrast is mainly manifested in backness and rounding and is also present in Norwegian, but absent in Polish, in which there is no near-high central rounded vowel. Finally, the Norwegian /i/—/ʏ/ contrast is mainly manifested in roundness and is absent in Polish and English, in which there are no front rounded vowels.

In the study we addressed the following research questions followed by associated predictions:

1. Will phonological contrasts be equally easy to detect and process in the native language (i.e., Polish) and non-native languages (i.e., English and Norwegian)?
   We predict the MMN effect to be larger in the native when compared with non-native vowel perception (Näätänen et al., 1997; Jakoby et al., 2011; Song and Iverson, 2018; Liang and Chen, 2022)
2. Will any significant distinctions emerge in L3/L*n* Norwegian as opposed to L2 English?
   The scale of the MMN effect in L2 when compared with L3/L*n* is difficult to predict due to the lack of previous studies which would focus on such a comparison. On the basis of previous L2 research, we can, however, tentatively assume that the MMN effect in L3/L*n* will be smaller relative to L1, and similar or smaller relative to L2. We can also predict the effect to be stronger in the more dominant and/or more proficient language.
3. What factors will play a crucial role in L2 and L3/L*n* phonological processing?
   Since studies in L2 phonemic perception point to the relevance of such factors as language proficiency (Liang and Chen, 2022), learning context (Peltola et al., 2003) and phonological aptitude (taken as a proxy indicator of the ability to discern between different sounds, see Díaz et al., 2016), we can also predict that these factors will affect the results of the current study on L2 as opposed to L3/L*n* phonemic perception.

2. Materials and methods

2.1. Participants

Twenty-one participants (mean age = 32.9, age range: 22–47, nine males) were recruited to take part in the study. They were all right-handed as assessed by the Edinburgh Handedness Inventory adapted from Oldfield (1971), with the mean laterality quotient (LQ) equal to 83.1% (range: 40.00%–100,00%, *SD* = 16.92%). All of the participants were originally from Poland and at the time of the study lived in Tromsø, Norway. Most of them were college graduates with an earned BA (*N* = 4), MA (*N* = 7) or PhD (*N* = 4) degree. Three participants were college students, and three reported high school as the highest completed level of education. According to self-reports, the ages of acquisition of the non-native languages was 9.48 years (range: 4–29, *SD* = 5.27) for L2 English and 27.33 years (range: 7[1]–43, *SD* = 8.21) for L3/L*n* Norwegian. For all the participants Polish was the only native language, and for all but two of them English was chronologically the first foreign language which they started learning at school or pre-school before puberty. The two participants started learning English at the ages of 15 (as the first foreign language) and of 29 (as the second foreign language, following Russian). The status of Norwegian differed more markedly among the participants: for various sub-groups, it was chronologically the third (*N* = 8), the fourth (*N* = 7), the fifth (*N* = 5), or even the sixth (*N* = 1) foreign language. The average length of residence in Norway equaled 7.79 years (range: 1–14, *SD* = 3.43).

The participants were asked to self-assess their knowledge of English and Norwegian in listening, speaking, reading and writing on a scale from 1 (very low) to 7 (proficient). In addition, their knowledge of the two investigated foreign languages was verified with the aid of two language proficiency tests taken immediately after the EEG session. The average score in the English proficiency test was 76.47% (range: 44.00–100.00%, *SD* = 15.85%), which would approximately correspond to the B2 level according to the CEFR proficiency scale. The average score in the Norwegian proficiency test was 58.65% (range: 22.22–94.44%, *SD* = 27.43%), which would approximately correspond to the A2 level according to the CEFR proficiency scale.

The summary of the participants' biographic details and language proficiency can be found in Table 1. A more detailed summary of the language history questionnaire as well as proficiency tests results for individual participants are included in the Supplementary materials.

1 There was only one participant who reported to be exposed to Norwegian before puberty, i.e., at the age of 7. The age range for the remaining 20 participants equals 20–43. It might be crucial that the early Norwegian learner has reported to be only passively exposed to Norwegian in childhood, and then started using the language at the age of 25 after moving to Norway.

None of the participants reported any neurological and psychiatric impairments nor any language-related issues (e.g., dyslexia, dysorthography). The participants signed an informed consent form before the experiment and received gift cards for their participation. Data from one participant (an L*n* speaker of Norwegian) was excluded from further analyzes due to technical issues.

2.2. Stimuli

Following Liang and Chen (2022), we used isolated vowels rather than vowels embedded in syllables within consonantal frameworks. Listeners are believed to process isolated vowels as speech thanks to the pre-attentive ability to extract the relevant F1/F2 formant ratio (Jakoby et al., 2011; Liang and Chen, 2022). Furthermore, using isolated vowels enabled us to investigate phonological contrast perception without any potential interference of co-articulation processes associated with syllable production, which are likely to be different in each of the three languages.

When it comes to the deviancy status of the selected vowels, in the Polish (L1) condition, the standard stimulus was the high central unrounded vowel /ɨ/ and the deviant stimulus was the high-mid front unrounded vowel /ɛ/ (as in the Polish words *byty* 'beingpl' and *bety* 'bed linenpl'). In the English (L2) condition, the standard stimulus was the near-high front unrounded vowel /ɪ/ and the deviant stimulus was the near-high central slightly rounded vowel /ʊ/ (as in the English words *fit* and *foot* respectively). In the Norwegian (L3/L*n*) condition, the standard stimulus was high front unrounded vowel /i/ and the deviant stimulus was the near-high front weakly rounded vowel /ʏ/ (as in the Norwegian words *sin* 'his$_{REFL}$' and *synd* 'shame' respectively). For the auditory stimuli, please visit our OSF repository: https://osf.io/2956a/?view_only=cf240fe1fab54b91a3aeab93c9e20423.

The vowels used in the current study were all synthesized with the aid of the PRAAT software (Boersma, 2001). Formant frequencies of Polish and English vowels were defined on the basis of the previous literature (Weckwerth and Balas, 2019 for Polish; Bjelaković, 2017 for English). Due to the lack of available literature, Norwegian vowels were generated based on the average values obtained from four native speakers of Norwegian (living in the Trondheim region). For all the synthesized stimuli the duration was 150 ms, the amplitude contour had a 3 ms linear onramp and 75 ms linear offramp, and the f0 trajectory had a steady linear fall from 140 Hz to 110 Hz. The formant values for each vowel as well as Euclidean distances between vowels used in the three language pairs are presented in Table 2. Our endeavors cannot be compared to the decisions made in previous studies, as their authors did not need to make choices concerning vowel pairs in three languages[2].

TABLE 1 The summary of the participants' biographic details and language proficiency.

Participants	
Biographic details	
Age	M = 32.9, range: 22–47, SD = 7.4
Gender	12 females, 9 males
Proficiency self-assessment	
L1 Polish	M = 6.94, range: 5.75–7, SD = 0.28
L2 English	M = 5.76, range: 4.5–7, SD = 0.91
L3/L*n* Norwegian	M = 3.74, range: 1–5.75, SD = 1.76
Proficiency tests results	
L2 English	M = 76.47%, SD = 15.85%
L3/L*n* Norwegian	M = 58.65%, SD = 27.43%

TABLE 2 Summary of vowel formant frequencies used for stimuli synthesis (in Hz) and Euclidean distances between vowels (in Hz and Bark).

Vowel	F1	F2	F3	F4
Polish /ɨ/	468	1948	2821	3425
Polish /ɛ/	675	1916	2722	3441
English /ɪ/	394	1828	2882	3409
English /ʊ/	390	1345	2896	3413
Norwegian /i/	357	1917	2587	3505
Norwegian /ʏ/	313	2015	2708	3549

Euclidean distance	/ɨ/: /ɛ/	/ɪ/: /ʊ/	/i/: /ʏ/
F1-F2 (Hz)	209	483	107
F1-F2 (Bark)	2.05	4.42	1.06
F1-F2-F3 (Hz)	232	483	161
F1-F2-F3 (Bark)	2.27	4.42	1.59

2.3. Procedure

The participants were tested individually in a sound-attenuated room. At the beginning of each session, they were asked to fill in a language history questionnaire (based on Li et al., 2020) and a survey concerning hand dominance based on the Edinburgh Handedness Inventory (Oldfield, 1971). During the EEG session, participants were seated comfortably while watching a muted cartoon (*Bolek and Lolek*) without subtitles. The choice of a cartoon over other genres was motivated by the necessity to use the most engaging visual material possible which would direct the subjects' attention away from the MMN-eliciting stimulus. Otherwise, attention-dependent ERP components might have overlapped with the MMN (Näätänen et al., 2007). Consequently, the participants were instructed to watch the movie carefully and attentively. They were also informed that they would be asked to answer a few questions about the content of the displayed story. The language of instruction was Polish.

The task sequence was controlled by a PC running Presentations software (Neurobehavioral Systems, http://www.neurobehavioralsystems.

2 Jakoby et al. (2011) tested two French stimuli among Hebrew-English bilinguals, and presented formant frequencies for French stimuli, Liang and Chen (2022) tested Chinese-English bilinguals on the perception of a standard that was claimed to exist in both Chinese Mandarin and English (but which had relatively low F2 values, denoting a back, rather than a central quality of /u/). Song and Iverson (2018) used sentences. Näätänen et al. (1997) tested the perception of Estonian and Finnish vowels: they used synthetic stimuli differing in F2 only.

com). The sounds were presented binaurally through in-ear headphones. The loudness of the stimuli was kept constant across all participants. Each trial began with a phonetic sound for 150 ms, followed by a silence of 700–1,000 ms. The phoneme pairs were presented in three separate language blocks (i.e., Polish, English and Norwegian), the order of which was counterbalanced across participants. In each language block, 600 standards and 60 deviants were presented at an intensity of 75 dB, with a probability of 90.9% and 9.1%, respectively. The standard/deviant ratio was in accordance with previous studies for which the deviant probability varied between 6.7% (Díaz et al., 2016) and 16.7% (Liang and Chen, 2022). Deviant stimuli appeared in a pseudorandomized order, with a minimum of three preceding standard stimuli. Each experimental block was followed by a short break of approximately 3 min, during which time no stimuli were presented, and the participants continued watching the movie in silence. After the EEG session, the participants were asked to complete a short test concerning the content of the movie they had watched. The test consisted of 10 multiple choice questions (e.g., "Where did the boys hide after they broke the glass in the window? in barrels/in the closet/in the chimney"). The main purpose of the test was to help us determine whether the participants remained focused while watching the movie and whether the pre-attentive state for listening was successfully created.

Further, the participants took part in a gating task conducted in English with the aim of determining the potential individual differences in terms of speech-specific capabilities in a foreign language. We selected English as the language of the task, given that it was chronologically the first and more advanced foreign language spoken by the participants (which was further confirmed by the results of the proficiency tests and self-reports). While designing the task, we adapted the procedure used by Sebastián-Gallés and Soto-Faraco (1999) and later by Sebastian-Galles and Baus (2005) who applied a two-alternative forced choice test. The participants' task was to identify the word whose fragment was presented via earphones by pressing "L" or "A" keys on the computer keyboard. The participants were also asked to assess how sure they were of their answer on a 7-point Likert scale. The experimental stimuli consisted of four monosyllabic word pairs including the /æ/−/ε/ contrast (i.e., *BAG-BEG*, *LAUGHED-LEFT*, *SHALL-SHELL*, *GAS-GUESS*). The alineation point (i.e., the point where the token words started to diverge) was determined on the basis of the visual inspection conducted with the aid of the PRAAT software (Boersma, 2001). This point was assumed to be "gate" 3. After the alineation point identification, the words were divided into other "gates" (i.e., fragments) by adding or subtracting 10 ms from the alineation point. Each member of the minimal pairs was presented two times, which resulted in 160 trials (4 pairs x 2 words x 10 "gates" x 2 presentations), with an optional break after 80 trials. The words were recorded by a native speaker of American English and presented at an intensity of 75 dB with the aid of the PsychoPy software (Peirce et al., 2019).

Finally, the participants were asked to complete two language proficiency tests: the Cambridge General English Assessment Test and the UiT Norwegian Placement Test. Thanks to this, we were able to adequately determine the participants' level of proficiency in both foreign languages. A single experimental session lasted about 2.5–3 h, including the EEG preparation, EEG recordings and all the remaining tasks.

All procedures were accepted by the Ethics Committee for Research with Human Participants at Adam Mickiewicz University.

2.4. EEG data acquisition and analysis

The EEG signal was recorded using Brain Products LiveAmp acquisition device at a 500 Hz sampling rate from 32 active electrodes placed at the elastic cap according to the extended 10–20 convention. The ground was positioned at AFz. In addition, two electrodes were placed at the outer canthus of each eye (HEOG1 and HEOG2) and two were placed below and above the right eye (VEOG1 and VEOG2). The signal was referenced online to FCz, and later re-referenced offline to the average of right and left mastoid bones (approximated from TP7 and TP8). Electrode impedances were kept below 10 kΩ. The EEG data was processed with the aid of the Brain Vision Analyzer 2 software (Brain Products, Gilching).

At the first preprocessing state, the data were filtered offline with a 0.1–30 Hz band-pass filter. Then, a semi-automatic ICA ocular correction was performed and the signals were re-referenced. Epochs time-locked to the onset of each stimulus were extracted between −200 to 800 ms. Only the standard stimuli which immediately preceded a deviant stimulus were considered in the analysis; hence, the number of standard events and the number of deviant events were equal in each language ($N = 60$). Baseline correction was performed in reference to pre-stimulus activity (i.e., −200 to 0 ms). The next step of the analysis involved the semi-automatic Raw Data Inspection (maximal allowed voltage step: 50 μV/ms, maximal allowed difference of values in intervals: 200 μV/ms, minimal allowed amplitude: 100 μV, maximal allowed amplitude: −100 μV). Epochs contaminated by ocular or muscular artifacts were rejected from further analysis, which resulted in the exclusion of 1.57% of trials (1.08% for Polish standards, 1.67% for Polish deviants, 1.42% for English standards, 1.25% for English deviants, 2.08% for Norwegian standards and 1.92% for Norwegian deviants).

The separately averaged waveforms for the standard and the deviant stimuli were computed for each subject and the difference waveforms were then created by subtracting the standard response from the response to the deviant stimulus. Following Luck and Gaspelin (2017), we first averaged the waveforms elicited by standard and deviant stimuli across all the language conditions and defined the time windows used in our analysis based on the collapsed waveforms. This approach revealed an increased negativity in the 100–200 ms time window, which was followed by a late negativity in the 350–800 ms time window. Since use of the 100–200 ms time window is in accordance with Kujala and Näätänen (2003) and the 350–800 ms time window was also previously used in the literature (Di Dona et al., 2022), we used these time windows to measure the effects in the three language conditions separately. The analyzed region of interest was the frontal-central brain area (F3, Fz, F4, FC1, FCz, FC2, C3, Cz, C4), given that both the MMN and LDN effects are typically observed in this scalp site (Ceponiene et al., 1998; Kujala and Näätänen, 2003).

The statistical analysis of the results was conducted with the aid of the R software (R Core Team, 2012). More specifically, we used the lme4 package (Bates et al., 2012) to perform a linear mixed effects analysis of the relationship between the processed language and the status of the processed sound as standard or deviant. The procedure was carried out twice: in the earlier time window (i.e., 100–200 ms) for the MMN effect and in the later time window (i.e., 350–800 ms) for the LDN effect. Language (i.e., Polish, English and Norwegian) and sound (i.e., Standard or Deviant) were included in the model as fixed effects. As random effects, we included intercepts for participants and electrodes. The model was applied to data averaged across 60 trials in each of the language and sound conditions. Visual inspection of the residual plots did not reveal any obvious deviations from homoscedasticity or normality in either of the two analyzed time window data sets. *p* -values were obtained by likelihood ratio tests of the full model with the interaction effect in question against the model with two main effects.

In the following step, we compared effect sizes for significant effects observed in the lme analysis. For this reason, we calculated the difference wave (i.e., deviant minus standard) for each participant, individually in each electrode. Once again, we used the lme4 package (Bates et al., 2012) to perform a linear mixed effects analysis of the relationship between the processed language and the scale of the MMN and LDN effects. The procedure was also repeated: in the earlier time window (i.e., 100–200 ms) for the MMN effect and in the later time window (i.e., 350–800 ms) for the LDN effect. Language (i.e., Polish, English and Norwegian) was included in the model as fixed effects. As random effects, we included intercepts for participants. Visual inspection of residual plots did not reveal any obvious deviations from homoscedasticity or normality in either of the two analyzed time window data sets. *p* -values were obtained by likelihood ratio tests of the full model with the main effect of Language against the model with no main effects.

3. Results

3.1. Comprehension test results

The overall results of the movie comprehension test were very high, with the average of 93.81% correct responses (range: 70–100%, SD = 8.05%). This means that the participants focused on the movie rather than the experimental stimuli which were processed pre-attentively.

3.2. ERP results

The analysis revealed the MMN effect elicited as a reaction to deviant sounds when compared with standard sounds. The component was particularly pronounced over frontal-central scalp sites and had a peak at around 150 ms after the sound onset. The MMN effect was followed by the LDN, with a peak around 450 ms after the sound onset, which was also particularly well visible over frontal-central scalp sites. Figure 1 below presents grand average ERPs elicited from nine representative (F3, Fz, F4, FC1, FCz, FC2, C3, Cz, C4) electrodes in response to standard sounds (dotted lines) and deviant sounds (solid line) in the three investigated languages. Figure 2 presents a voltage difference map (deviant minus standard) in the analyzed time windows, i.e., 100–200 ms (for MMN) and 350–800 ms (for LDN).

Descriptive statistics for sound and language conditions in the two time windows of interest are presented in Table 3. Figures displaying mean amplitude values observed in each condition and each target language as well as mean amplitude differences in each target language are included in the Supplementary material.

3.2.1. MMN

In the 100–200 ms time window, model comparison revealed a statistically significant interaction effect of language and sound ($\chi^2(2) = 21.554$; $p < 0.001$). To further examine the significant interaction effect, Tukey based pairwise comparisons were performed, which revealed that in each language deviant sounds elicited significantly more negative amplitudes than standard sounds ($p < 0.001$) (see Table 4).

Since we observed statistically significant differences in each analyzed language, we then calculated the difference wave (i.e., deviant minus standard) for each participant, individually in each electrode, and conducted a linear mixed effects analysis of the relationship between the processed language and the scale of the MMN effect (recall Section 2.4 for details). Descriptive statistics for language conditions are presented in Table 5. The deviant minus standard difference was the greatest in L1 Polish, a bit smaller in L2 English and the smallest in L3/L*n* Norwegian. In the 100–200 ms time window, model comparison revealed a statistically significant main effect of language ($\chi^2(2) = 28.505$; $p < 0.001$). Tukey based pairwise comparisons (see Table 6) revealed that the deviant minus standard difference was significantly higher in L1 Polish than in L3/L*n* Norwegian (Estimate = 0.775, $p < 0.001$) and significantly higher in L2 English than in L3/L*n* Norwegian (Estimate = 0.440, $p < 0.01$). The difference between L1 Polish and L2 English, however, was not statistically significant (Estimate = 0.336, $p = 0.0521$).

3.2.2. LDN

The statistical analysis conducted in the 350–800 ms time window revealed a statistically significant language and sound interaction ($\chi^2(2) = 12.36$; $p < 0.01$). Tukey based pairwise comparisons revealed that in each language deviant sounds elicited significantly more negative amplitudes than standard sounds ($p < 0.001$).

As in the case of the 100–200 ms time window, we observed a statistically significant negativity associated with the occurrence of a deviant in each investigated language. Consequently, we conducted an additional linear mixed effect analysis based on a model which included the deviant minus standard difference as a dependent variable. In the 350–800 ms time window this kind of analysis also revealed a statistically significant main effect of language ($\chi^2(2) = 16.75$; $p < 0.001$). Tukey based pairwise comparisons showed a statistically significant difference between L1 Polish and L2 English as well as between L1 Polish and L3/L*n* Norwegian, with the LDN effect significantly stronger in the case of Polish (Estimate = 0.617, $p < 0.001$ and Estimate = 0.522, $p < 0.01$ respectively). The difference between L2 English and L3/L*n* Norwegian was not statistically significant (Estimate = 0.095, $p = 0.825$).

3.3. Gating task results

In terms of the gating task, we calculated the mean accuracy for each participant as well as the mean 'gate' at which the words were recognized. While calculating the mean accuracy, we only took into account the answers which satisfied the following two criteria: (a) the decision concerning the selected word was not changed afterwards

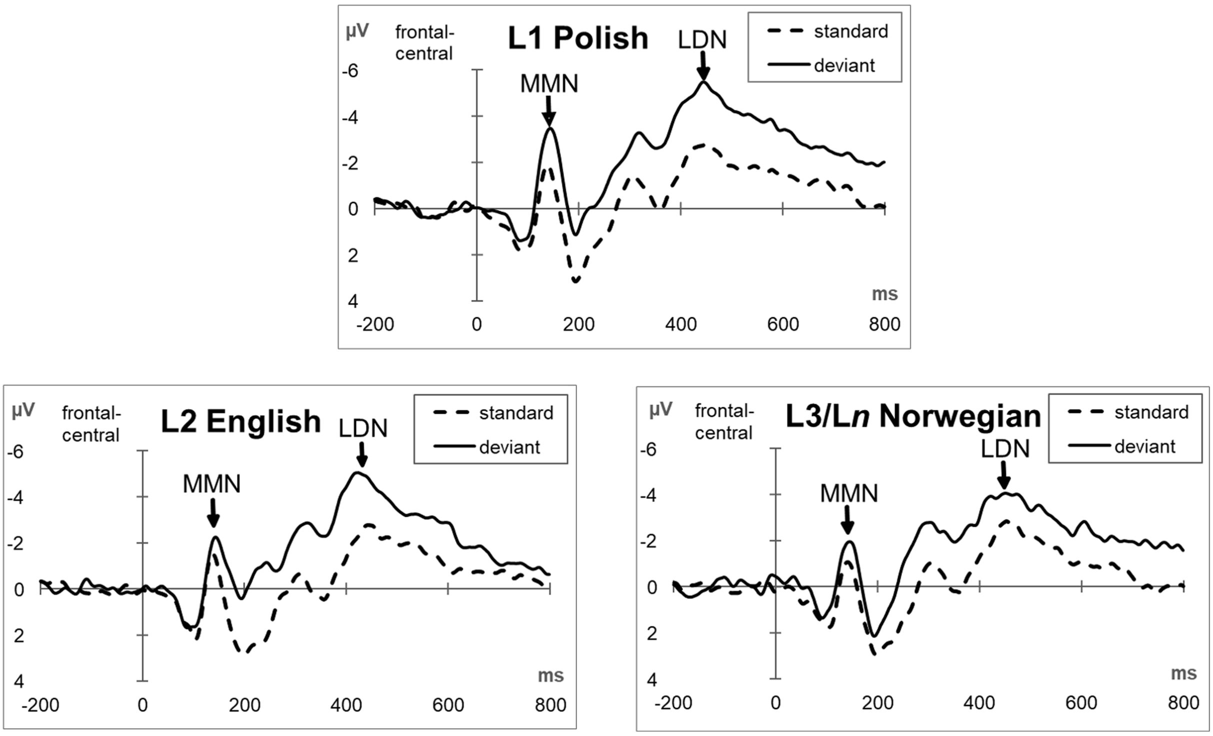


FIGURE 1
The grand average ERPs time-locked to the onset of the phoneme for the standard (dotted line) and deviant stimuli (solid line) elicited from nine representative electrodes (F3, Fz, F4, FC1, FCz, FC2, C3, Cz, C4) in the three investigated languages.

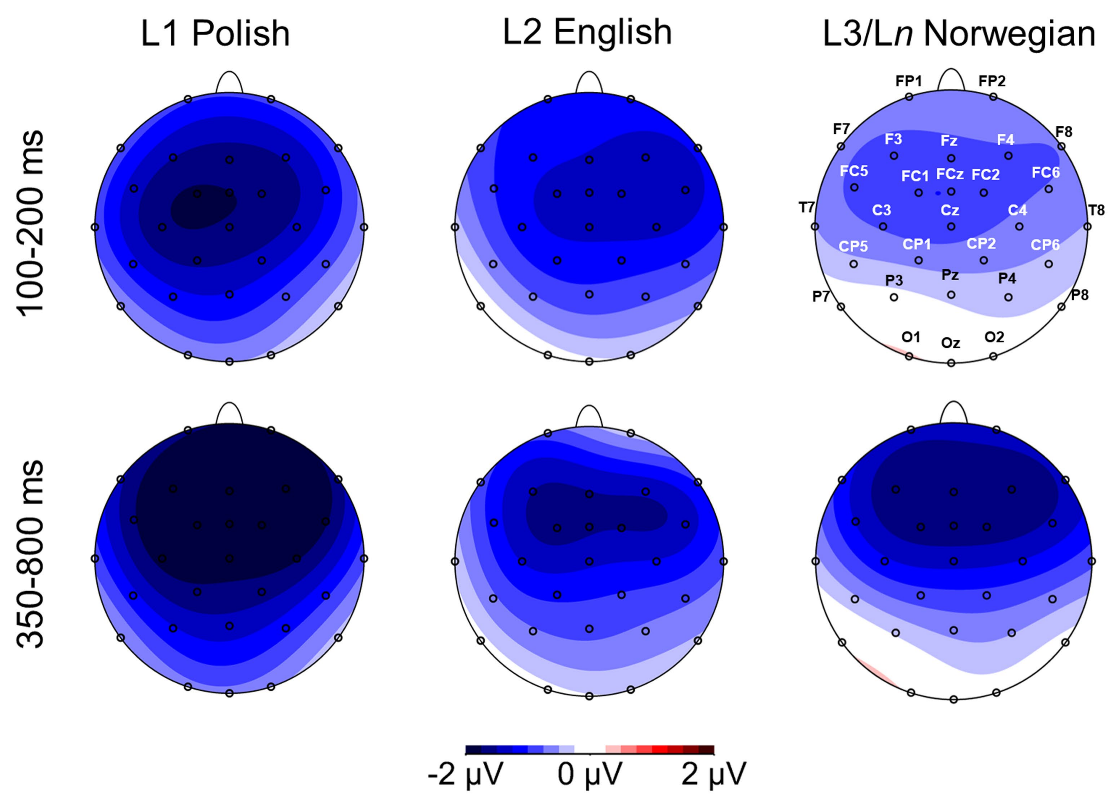


FIGURE 2
Topographic distribution of voltage differences between deviant and standard conditions in the three investigated languages in the 100-200 and 350-800 ms time windows.

TABLE 3 Descriptive statistics for the experimental conditions: standard/deviant and Polish/English/Norwegian.

	Emmean	SE	df	Lower. CL	Upper. CL
Time window: 100–200ms					
L1 Polish					
Standard	0.47	0.381	23.3	−0.32	1.26
Deviant	−1.17	0.381	23.3	−1.95	−0.38
L2 English					
Standard	0.81	0.381	23.3	0.03	1.60
Deviant	−0.48	0.381	23.3	−1.27	0.31
L3/*Ln* Norwegian					
Standard	0.78	0.381	23.3	−0.01	1.57
Deviant	−0.07	0.381	23.3	−0.86	0.72
Time window: 350–800ms					
L1 Polish					
Standard	−1.36	0.28	28	−1.93	−0.79
Deviant	−3.42	0.28	28	−3.99	−2.85
L2 English					
Standard	−1.20	0.28	28	−1.77	−0.63
Deviant	−2.64	0.28	28	−3.21	−2.07
L3/*Ln* Norwegian					
Standard	−1.11	0.28	28	−1.68	−0.54
Deviant	−2.64	0.28	28	−3.21	−2.08

TABLE 4 Statistically significant pairwise comparisons between experimental conditions: standard/deviant and Polish/English/Norwegian.

Compared conditions	Estimate	SE	df	t.ratio	p-value
Time window: 100–200ms					
L1 Polish					
contrast: deviant – standard	−1.63	0.119	1054	−13.755	<0.0001
L2 English					
contrast: deviant – standard	−1.30	0.118	1054	−10.938	<0.0001
L3/*Ln* Norwegian					
contrast: deviant – standard	−0.86	0.119	1054	−7.176	<0.0001
Time window: 350–800ms					
L1 Polish					
contrast: deviant – standard	−2.06	0.133	1057	−15.424	<0.0001
L2 English					
contrast: deviant – standard	−1.44	0.133	1057	−10.802	<0.0001
L3/*Ln* Norwegian					
contrast: deviant – standard	−1.54	0.133	1057	−11.517	<0.0001

Degrees-of-freedom method: Kenward-Roger. *p*-value adjustment: Tukey method for comparing a family of 3 estimates.

and (b) the level of confidence was assessed as at least 4 in a 7-point Likert scale. On average, the accuracy score equaled 78.87% (range: 50.00–100%, $SD = 14.72\%$) and the words were recognized correctly after the eighth 'gate' ($M = 8.23$, range: 6.4–10, $SD = 1.08$).

To check whether the participants' phonological aptitude (indexed by the results of the gating task) would influence the MMN or the LDN effect in English or Norwegian, we conducted several simple linear regression analyzes. None of them, however, yielded statistically significant results. They all included the overall gating accuracy or mean gates at which the words were recognized as predictor variables. The deviant minus standard value obtained for each participant in the respective language conditions (i.e., English or Norwegian) and time windows (i.e., 100–200 ms or 350–800 ms) were included as response variables. However, the amplitude of the investigated components was not significantly predicted by the participants' gating accuracy (MMN in the English condition: $p = 0.629$, $R^2 = 0.013$, LDN in the English condition: $p = 0.949$, $R^2 < 0.001$; MMN in the Norwegian condition: $p = 0.090$, $R^2 = 0.151$; LDN in the Norwegian condition: $p = 0.969$, $R^2 < 0.001$) nor the mean of the "gates" at which the words were recognized (MMN in the English condition: $p = 0.394$, $R^2 = 0.041$; LDN in the English condition: $p = 0.870$, $R^2 < 0.010$; MMN in the Norwegian condition: $p = 0.939$, $R^2 < 0.001$; LDN in the Norwegian condition: $p = 0.114$, $R^2 = 0.133$).

Further, to verify whether the participants' proficiency (self-assessed and further verified by two foreign language tests; recall Table 1), dominance (associated with the frequency of language use) or age of acquisition (self-reported in the LHQ) would affect the scale of the MMN or the LDN effect, we conducted additional linear mixed effect analyzes, independently for MMN and LDN. The first two analyzes included the self-reported proficiency scores as predictor variables and the deviant minus standard value obtained for each participant in the respective language conditions as a criterion variable. As random effects, we included intercepts for participants. The proficiency score was revealed to predict the scale of the MMN effect ($p < 0.001$, $R^2 = 0.227$) but no statistically significant result was obtained in the case of the LDN ($p = 0.153$, $R^2 = 0.405$). Further, the MMN and LDN amplitudes were both significantly predicted by the participants' dominance operationalized in terms of the number of hours per week which they reported in the LHQ (MMN: $p < 0.001$, $R^2 = 0.266$; LDN: $p < 0.023$, $R^2 = 0.413$). Finally, a correlation was found between the scale of both ERP effects and the participants' age of acquisition (MMN: $p < 0.001$, $R^2 = 0.276$; p, LDN: $p < 0.001$, $R^2 = 0.422$).

4. Discussion

The main objective of the current study was to shed more light on non-native phonological contrast perception – a phenomenon particularly relevant nowadays, with multilingualism having already become a norm in the modern globalized world (e.g., Aronin and Singleton, 2008). Previous research has demonstrated that the processing of phonological contrasts is typically hampered in non-native when compared with native languages (Jakoby et al., 2011; Song and Iverson, 2018; Liang and Chen, 2022). In the present work, we aimed to extend the scope of research in the field so that it involved two non-native languages. This way, we hoped to contribute to the

TABLE 5 Descriptive statistics for the MMN effect expressed in terms of the deviant minus standard difference in the three language conditions: Polish, English and Norwegian.

	Emmean	SE	df	Lower. CL	Upper. CL
Time window: 100–200ms					
L1 Polish	−1.63	0.191	28.8	−2.02	−1.24
L2 English	−1.30	0.191	28.8	−1.69	−0.91
L3/L*n* Norwegian	−0.86	0.191	28.8	−1.25	−0.45
Time window: 350–800ms					
L1 Polish	−2.06	0.309	23	−2.70	−1.42
L2 English	−1.44	0.309	23	−2.08	−0.8
L3/L*n* Norwegian	−1.54	0.309	23	−2.18	−0.898

TABLE 6 Pairwise comparisons for the MMN effect expressed in terms of the deviant minus standard difference in the three language conditions: Polish, English and Norwegian.

Compared conditions	Estimate	SE	df	t.ratio	*p*-value
Time window: 100–200ms					
Contrast: English – Polish	0.336	0.144	518	2.334	0.0521
Contrast: Norwegian – Polish	0.775	0.144	518	5.387	<0.0001
Contrast: English – Norwegian	−0.440	0.144	518	−3.053	0.0067
Time window: 350–800ms					
Contrast: English – Polish	0.617	0.161	518	3.824	<0.001
Contrast: Norwegian – Polish	0.522	0.161	518	3.233	<0.01
Contrast: English – Norwegian	0.095	0.161	518	0.591	0.825

Degrees-of-freedom method: Kenward-Roger; *p*-value adjustment: Tukey method for comparing a family of 3 estimates.

ongoing discussion on the perception of native as opposed to non-native phonemes by multilingual speakers (Cabrelli and Wrembel, 2016). Specifically, we tested vowel contrast perception among L1 Polish-L2 English-L3/L*n* Norwegian speakers.

The first research question investigated whether phonological contrasts would be equally easy to detect and process in the native language (i.e., Polish) and in non-native languages (i.e., English and Norwegian). Following previous authors, we predicted that the MMN response would be stronger in native vowel perception when compared with non-native vowel perception (Näätänen et al., 1997; Jakoby et al., 2011; Song and Iverson, 2018; Liang and Chen, 2022). This hypothesis, however, was confirmed only in the case of L3/L*n* Norwegian when compared with L1 Polish. While each vowel contrast elicited a statistically significant MMN effect (Table 4), there was no statistically significant difference between the effect observed for L1 Polish and the effect elicited in L2 English (Table 6). This finding suggests that – perhaps with sufficient exposure – phonological perception mechanisms might be equally developed in the non-native language when compared with native language. Such a result is also, at least partly, in accordance with the study of Winkler et al. (1999), who found a similar MMN response to Finnish vowel contrasts in native speakers of Finnish and in naturalistic late learners of Finnish. Very importantly, however, the MMN effect observed in the current study for L3/L*n* Norwegian was statistically weaker when compared with L1 Polish. This confirms that, even for foreign languages acquired in a naturalistic setting, phonological contrasts may not always be detected as easily as in the case of one's mother tongue.

The second research question focused on the possible emergence of any significant distinctions between L3/L*n* and L2 English. We predicted that the effect would be stronger in the more dominant and/or more proficient language. Our findings show statistically significant differences between the two foreign languages: the MMN effect was significantly stronger in L2 English when compared with L3/L*n* Norwegian. This is in accordance with our hypothesis that the effect would be enhanced for the more dominant and/or proficient foreign language. As indicated by the results of language proficiency tests, the participants in the current study – despite living in Norway – were much more proficient in English than in Norwegian. On average, they obtained 76.47% in the English proficiency test as opposed to 58.65% in the Norwegian proficiency test, and the outcomes were further supported self-assessment ratings (5.76 as opposed to 3.74 respectively). What is more, English has also turned out to be the foreign language which was more frequently used by the participants (mostly in the international work environment). Out of the 20 speakers whose data was included in the final analysis, 10 reported using English most frequently out of the three investigated languages, seven used Polish most frequently, three used English and Polish to a similar degree, but only one indicated Norwegian as their most frequently used language.

This observation is closely related to the third research question which explored the factors that might play a crucial role in L2 and L3/L*n* processing. As space does not allow for the consideration of every single one of these factors, we preliminarily distinguished AoA, proficiency, dominance and phonological aptitude as potential predictors of successful phoneme discrimination in the two non-native languages. We sought to determine whether any of these factors (measured by additional tests and self-reports) would influence the degree of the investigated ERP effects. Indeed, we found out that AoA, proficiency and language dominance impacted the MMN effect, and AoA and language dominance affected the LDN effect.In fact, the more global processing patterns reflected in the differences between the investigated language pairs (i.e., L1 vs. L2, L1 vs. L3/L*n*, L2 vs. L3/L*n*) might also enable us to point to language dominance and proficiency as two factors which seem to be of particular relevance in mastering the discrimination of non-native phonemes. This also remains in accordance with previous research on phonological discrimination mechanisms in L2 (Jakoby et al., 2011; Archila-Suerte et al., 2012; Liang and Chen, 2022). However, since the results of the current study cannot fully disentangle the effects of proficiency and dominance (as the participants were apparently both more dominant and more proficient in English than in Norwegian), this distinction should be further investigated in the future.

What is also noteworthy is that the vast majority of participants started learning English in their early childhood (on average, at around the age of nine) and acquired Norwegian much later in life, well after puberty (i.e., at around 27 years of age). The measure of success in second language learning, and especially in terms of pronunciation, is frequently associated with the speaker's age of acquisition/arrival. For example, several linguistic studies observed a positive correlation between the age of arrival to the country in which the target language is spoken and the perceived strength of accentedness (see Bongaerts et al., 1995; Flege et al., 1995, 1999, among many others). This correlation also seems to be corroborated by the current study results as reflected in the MMN difference between L2 English and L3/L*n* Norwegian.

In addition to the MMN component, deviant stimuli in all three languages have also elicited the LDN response, a component whose functional significance still remains largely unsettled. Some authors have postulated that this component reflects the pre-attentive cognitive evaluation of the stimulus, while others have associated it with the extraction of the phonological difference between the standard and deviant stimuli, the reorientation of attention to the original task, or the formation of new phonological representations (see Jakoby et al., 2011 for a discussion). In the context of non-native phoneme perception, the LDN was larger in successful compared to unsuccessful language learners (Jakoby et al., 2011) and in more advanced compared to elementary ones (Liang and Chen, 2022). These findings seem to support the last explanation proposed above, i.e., that the LDN might index a successful formation of memory traces associated with specific phonemic representations – an explanation proposed also by Barry et al. (2009). In the current study, the LDN was largest in L1 Polish, smaller in L3/L*n* Norwegian, and the smallest in L2 English, with the difference between L2 English and L3/L*n* Norwegian not statistically significant. Quite crucially, the difference between the non-native languages in question reached the level of statistical significance in the MMN time window. When interpreted together, these two findings might be viewed as tentatively supporting the idea that the LDN if functionally independent from the MMN as well as the claim that the component indexes the formation of new phonological representations (in this case, in the non-native languages). These hypotheses would need to be further verified by a longitudinal study examining the yet to established functional role of the LDN over a longer period of time.

One limitation of the current research is that – as many studies focused on multilingual language processing – it used a relatively small sample size (i.e., 20 trilingual participants). What is more, the experiment could have ideally used a mirror design, e.g., L1 Polish ->L2 English ->L3/L*n* Norwegian vs. L1 Polish ->L2 Norwegian ->L3/L*n* English (see Puig-Mayenco et al., 2020 for a discussion). Such a solution would enable us to directly compare the influence of language status on pre-attentive phonological processing and eliminate the potential confounds associated with the processing of specific vowel contrasts selected for each investigated language system. However, it would be extremely hard to find such a mirror group due to the prevalence of English as an L2 at the early stages of education. Possibly, a different combination of languages could be used in future research. In similar vein, the phonological aptitude test should ideally measure phoneme discrimination abilities in all three languages under investigation, i.e., not only in L2 English but also in L1 Polish and L3/L*n* Norwegian.

To the best of our knowledge, the current experiment was the first passive-oddball study to involve multilingual listeners. It resulted in several novel findings concerning multilingual phonological processing. Most crucially, the analysis of the ERP results revealed that the MMN was modulated by language. The MMN response in L3/*Ln* Norwegian was smaller when compared with L2 English and L1 Polish. At the same time, the LDN response in both L2 English and L3/*Ln* Norwegian was smaller when compared with L1 Polish. This provides preliminary, yet clear evidence that the foreign language status modulates auditory language processing. Living in an L3 environment does not then seem to be a guarantee of the development of native-like phonemic discrimination. Rather, it is language dominance, proficiency and age of acquisition which seem to be the most vital predictors of successful phonological difference extraction as well as the subsequent formation of new phonological representations.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary material, further inquiries can be directed to the corresponding author.

Ethics statement

The studies involving humans were approved by Ethics Committee for Research with Human Participants at Adam Mickiewicz University. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

HK: Conceptualization, Formal analysis, Investigation, Methodology, Visualization, Writing – original draft. KR: Conceptualization, Formal analysis, Methodology, Supervision, Visualization, Writing – review & editing. AB: Conceptualization, Methodology, Writing – review & editing. ZC: Conceptualization, Investigation, Writing – review & editing. CC: Investigation, Writing – review & editing. MW: Conceptualization, Funding acquisition, Methodology, Project administration, Supervision, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. The research leading to these results has received funding from the Norwegian Financial Mechanism 2014–2021 project number 2019/34/H/HS2/00495.

Acknowledgments

We would like to thank Jason Rothman and Vincent DeLuca for the opportunity to conduct the study in the PoLaR lab at UiT. We are also grateful to Kamil Kaźmierski for his comments on data analysis and to Anna Skałba for her assistance with the EEG recordings.

Conflict of interest

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

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2023.1270743/full#supplementary-material

References

Alho, K., Huotilainen, M., and Näätänen, R. (1995). Are memory traces for simple and complex sounds located in different regions of auditory cortex? Recent MEG studies. *Electroencephalogr. Clin. Neurophysiol.* 44, 197–203.

Archila-Suerte, P., Zevin, J., Bunta, F., and Hernandez, A. E. (2012). Age of acquisition and proficiency in a second language independently influence the perception of non-native speech. *Bilingualism Lang. Cogn.* 15, 190–201. doi: 10.1017/S1366728911000125

Aronin, L., and Singleton, D. (2008). Multilingualism as a new linguistic dispensation. *Int. J. Multiling.* 5, 1–16. doi: 10.2167/ijm072.0

Barry, J. G., Hardiman, M. J., and Bishop, D. V. (2009). Mismatch response to polysyllabic nonwords: a neurophysiological signature of language learning capacity. *PLoS One.* 4:e6270. doi: 10.1371/journal.pone.0006270

Bates, D. M., Maechler, M., and Bolker, B. (2012). *lme4: linear mixed-effects models using S4 classes. R package version 0.999999-0.*

Bjelaković, A. (2017). The vowels of contemporary RP: vowel formant measurements for BBC newsreaders. *English Lang. Linguistics* 21, 501–532. doi: 10.1017/S1360674316000253

Boersma, P. (2001). Praat, a system for doing phonetics by computer. *Glot. Int.* 5, 341–345.

Bongaerts, T., Planken, B., and Schils, E. (1995). "Can late starters attain a native accent in a foreign language? A test of the critical period hypothesis" in *The age factor in second language acquisition*. eds. D. Singleton and Z. Lengyel (Clevedon, GB: Multilingual Matters Limited), 30–50.

Cabrelli, A. J., and Wrembel, M. (2016). Investigating the acquisition of phonology in a third language – a state of the science and an outlook for the future. *Int. J. Multiling.* 13, 395–409. doi: 10.1080/14790718.2016.1217601

Ceponiene, R., Cheour, M., and Näätänen, R. (1998). Interstimulus interval and auditory event-related potentials in children: evidence for multiple generators. *Electroencephalogr. Clin. Neurophysiol.* 108, 345–354. doi: 10.1016/S0168-5597(97)00081-6

Cohen, M. X. (2014). *Analyzing neural time series data: Theory and practice*. Cambridge, The MIT Press.

Di Dona, G., Scaltritti, M., and Sulpizio, S. (2022). Formant-invariant voice and pitch representations are pre-attentively formed from constantly varying speech and non-speech stimuli. *Eur. J. Neurosci.* 56, 4086–4106. doi: 10.1111/ejn.15730

Díaz, B., Mitterer, H., Broersma, M., Escara, C., and Sebastián-Gallés, N. (2016). Variability in L2 phonemic learning originates from speech-specific capabilities: an MMN study on late bilinguals. *Bilingual. Lang. Cogn.* 19, 955–970. doi: 10.1017/S1366728915000450

Dziubalska-Kołaczyk, K., and Wrembel, M. (2022). "Natural growth theory of acquisition (NGTA): evidence from (mor)phonotactics" in *Theoretical and practical developments in English speech assessment, research, and training. Second language learning and teaching*. eds. V. G. Sardegna and A. Jarosz (New York: Springer), 281–298.

Flege, J. E., Munro, M. J., and MacKay, I. R. A. (1995). Factors affecting strength of perceived foreign accent in a second language. *J. Acoust.* 97, 3125–3134. doi: 10.1121/1.413041

Flege, J. E., Yeni-Komshian, G. H., and Liu, S. (1999). Age constraints on second-language acquisition. *J. Mem. Lang.* 41, 78–104. doi: 10.1006/jmla.1999.2638

Hawkins, S., and Midgley, J. (2005). Formant frequencies of RP monophthongs in four age groups of speakers. *J. Int. Phon. Assoc.* 35, 183–199. doi: 10.1017/S0025100305002124

Jakoby, H., Goldstein, A., and Faust, M. (2011). Electrophysiological correlates of speech perception mechanisms and individual differences in second language attainment. *Psychophysiology* 48, 1516–1530. doi: 10.1111/j.1469-8986.2011.01227.x

Jassem, W. (2003). Polish. *J. Int. Phon. Assoc.* 33, 103–107. doi: 10.1017/S0025100303001191

Kaan, E. (2007). Event-related potentials and language processing: a brief overview. *Lang. Linguist. Compass* 1, 571–591. doi: 10.1111/j.1749-818X.2007.00037.x

Kristoffersen, G. (2000). *The phonology of Norwegian*. Oxford: Oxford University Press.

Kujala, A., and Näätänen, R. (2003). "Auditory environment and change detection as Iindexed by the mismatch negativity (MMN)" in *Detection of change*. ed. J. Polich (Boston, MA: Springer)

Li, P., Zhang, F., Yu, A., and Zhao, X. (2020). Language history questionnaire (LHQ3): an enhanced tool for assessing multilingual experience. *Bilingual. Lang. Cogn.* 23, 938–944. doi: 10.1017/S1366728918001153

Liang, L., and Chen, B. (2022). The non-native phonetic perception mechanism utilized by bilinguals with different L2 proficiency levels. *Int. J. Biling.* 26, 368–386. doi: 10.1177/13670069211058275

Luck, S. J. (2005). *An introduction to the event-related potential technique*. Cambridge, MA: MIT Press.

Luck, S. J., and Gaspelin, N. (2017). How to get statistically significant effects in any ERP experiment (and why you shouldn't). *Psychophysiology* 54, 146–157. doi: 10.1111/psyp.12639

Maddieson, I. (2013). Front rounded vowels. *WALS Online (v2020.3)*, eds M. S. Dryer and M. Z. Haspelmath. Available at: http://wals.info/chapter/11 (Accessed July 13, 2023).

Näätänen, R., Lehtokoski, A., Lennes, M., Cheour, M., Houtilainen, M., Iivonen, A., et al. (1997). Language-specific phoneme representations revealed by electric and magnetic brain responses. *Nature* 385, 432–434. doi: 10.1038/385432a0

Näätänen, R., Paavilainen, P., Rinne, T., and Alho, K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing: a review. *Clin. Neurophysiol.* 118, 2544–2590. doi: 10.1016/j.clinph.2007.04.026

Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* 9, 97–113. doi: 10.1016/0028-3932(71)90067-4

Peirce, J. W., Gray, J. R., Simpson, S., MacAskill, M. R., Höchenberger, R., Sogo, H., et al. (2019). PsychoPy2: experiments in behavior made easy. *Behav. Res. Methods* 51, 195–203. doi: 10.3758/s13428-018-01193-y

Peltola, M. S., Kujala, T., Toumainen, J., Ek, M., Aaltonen, O., and Näätänen, R. (2003). Native and foreign vowel discrimination as indexed by the mismatch negativity (MMN) response. *Neurosci. Lett.* 352, 25–28. doi: 10.1016/j.neulet.2003.08.013

Peltola, M. S., Tamminen, H., Toivonen, H., Kujala, T., and Näätänen, R. (2012). Different kinds of bilinguals – different kinds of brains: the neural organisation of two languages in one brain. *Brain Lang.* 121, 261–266. doi: 10.1016/j.bandl.2012.03.007

Polich, J. (2012). "Neuropsychology of P300" in *The Oxford handbook of event-related potential components*. eds. S. Luck and E. S. Kappenman (Oxford: Oxford University Press), 159–188.

Puig-Mayenco, E., González Alonso, J., and Rothman, J. (2020). A systematic review of transfer studies in third language acquisition. *Second. Lang. Res.* 36, 31–64. doi: 10.1177/0267658318809147

R Core Team (2012). *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing.

Roehm, D., Bornkessel-Schlesewsky, I., Rösler, F., and Schlesewsky, M. (2007). To predict or not to predict: influences of task and strategy on the processing of semantic relations. *J. Cogn. Neurosci.* 19, 1259–1274. doi: 10.1162/jocn.2007.19.8.1259

Sebastian-Galles, N., and Baus, C. (2005). “On the relationship between perception and production in L2 categories” in *Twenty-first century psycholinguistics: Four cornerstones*. ed. A. Cutler (London: Routledge), 279–292.

Sebastián-Gallés, N., and Soto-Faraco, S. (1999). Online processing of native and non-native phonemic contrasts in early bilinguals. *Cognition* 72, 111–123. doi: 10.1016/S0010-0277(99)00024-4

Slabakova, R. (2017). The scalpel model of third language acquisition. *Int. J. Biling.* 21, 651–665. doi: 10.1177/1367006916655413

Song, J., and Iverson, P. (2018). Listening effort during speech perception enhances auditory and lexical processing for non-native listeners and accents. *Cognition* 179, 163–170. doi: 10.1016/j.cognition.2018.06.001

Weckwerth, J., and Balas, A. (2019). “Selected aspects of Polish vowel formants” in *Approaches to the study of sound structure and speech*. eds. M. Wrembel, A. Kiełkiewicz-Janowiak and P. Gąsiorowski (London: Routledge), 338–348.

Westergaard, M., Mitrofanova, N., Mykhaylyk, R., and Rodina, Y. (2017). Crosslinguistic influence in the acquisition of a third language: the linguistic proximity model. *Int. J. Biling.* 21, 666–682. doi: 10.1177/1367006916648859

Winkler, I., Kujala, T., Tiitinen, H., Sivonen, P., Alku, P., Lehtokoski, A., et al. (1999). Brain responses reveal the learning of foreign language phonemes. *Psychophysiology* 36, 638–642. doi: 10.1111/1469-8986.3650638

Wottawa, J., Adda-Decker, M., and Isel, F. (2022). Neurophysiology of non-native sound discrimination: evidence from German vowels and consonants in successive French–German bilinguals using an MMN oddball paradigm. *Bilingual. Lang. Cogn.* 25, 137–147. doi: 10.1017/S1366728921000468

Wrembel, M. (2015). *In search of a new perspective: Cross-linguistic influence in the acquisition of third language phonology*. Poznań: Wydawnictwo Naukowe UAM.

Wrembel, M., and Cabrelli Amaro, J. (eds.) (2018). *Advances in the investigation of L3 phonological acquisition*. London: Routledge.

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 30 November 2023
DOI 10.3389/fpsyg.2023.1293529

Check for updates

OPEN ACCESS

EDITED BY
Monika M. Polczynska,
University of California, Los Angeles,
United States

REVIEWED BY
Alena Stasenko,
University of California San Diego,
United States
Xiaohua Cao,
Zhejiang Normal University, China

*CORRESPONDENCE
Kichun Nam
✉ kichun@korea.ac.kr

RECEIVED 13 September 2023
ACCEPTED 09 November 2023
PUBLISHED 30 November 2023

CITATION
Kim S and Nam K (2023) Decoding foveal word recognition: the role of interhemispheric inhibition in bilateral hemispheric processing. *Front. Psychol.* 14:1293529.
doi: 10.3389/fpsyg.2023.1293529

Decoding foveal word recognition: the role of interhemispheric inhibition in bilateral hemispheric processing

Sangyub Kim[1] and Kichun Nam[2]*

[1]Wisdom Science Center, Korea University, Seoul, Republic of Korea, [2]School of Psychology, Korea University, Seoul, Republic of Korea

Extant research has largely favored the Split Fovea Theory (SFT) over the Bilateral Projection Theory (BPT) in the context of foveal word recognition. SFT posits that during foveal fixation, letters in the left and right visual fields are projected to their respective contralateral hemispheres, thereby facilitating a division of labor across the bilateral hemispheres. This division may serve as a regulatory mechanism to mitigate redundant processing in both hemispheres. The present investigation conducted two experiments utilizing Korean visual words to explore whether this hemispheric division in foveal word recognition is a strategy to circumvent potential interhemispheric inhibition arising from duplicated processing. Experiment 1 established the suitability of Korean visual words for studies involving both unilateral and bilateral presentations. Experiment 2 revealed that the split presentation of a word elicited greater accuracy compared to its identical presentation in the bilateral visual fields. These findings lend credence to the notion that interhemispheric inhibition may drive the hemispheres to engage in divided labor, thereby reducing processing redundancy in foveal word recognition.

KEYWORDS

hemispheric coordination, interhemispheric inhibition, split fovea theory, bilateral projection theory, foveal word recognition

1 Introduction

Building upon the foundational insights of Hellige (1993), the intricate interplay between the left and right cerebral hemispheres in cognitive functioning has been substantiated. Prior empirical investigations delineate that these hemispheres operate in a parallel yet autonomous fashion, each serving as a discrete computational entity (Iacoboni and Zaidel, 1996; Lindell et al., 2007). Within the neural architecture, a dynamically adaptive network—both functionally and structurally—facilitates a blend of concurrent and sequential processing modalities, thereby enhancing the computational efficiency of each hemisphere. Despite the myriad cognitive advantages conferred by this hemispheric specialization, an inherent regulatory mechanism within the inter-hemispheric interface fosters collaborative interactions. Specifically, inhibitory modulatory processes serve to integrate and harmonize the outputs emanating from each hemisphere, thereby precluding the emergence of potential computational discord between them.

Within the specialized domain of foveal vision as it pertains to visual word recognition, the academic landscape has been characterized by a dichotomy of theoretical paradigms (e.g., Ellis and Brysbaert, 2010). The first of these, the Split Fovea Theory (SFT), posits that the visual stimuli corresponding to the left segment of a word—based on the point of fixation—are selectively projected onto the right cerebral hemisphere, and conversely, the right segment is projected onto the left hemisphere. This segregated information subsequently undergoes interhemispheric transfer, primarily facilitated through neural conduits such as the corpus callosum (Brysbaert, 2004; Lavidor and Walsh, 2004). In contrast, the Bilateral Projection Theory (BPT) contends that foveally presented words are simultaneously propagated to both hemispheres, reserving contralateral projection exclusively for parafoveal words (Bunt et al., 1977). While both theories offer explanatory frameworks for the bilateral hemispheric mechanisms underlying foveal word recognition, recent empirical inquiries have increasingly lent credence to the SFT model (e.g., Brysbaert et al., 1996; Portin et al., 1998; Lavidor and Walsh, 2004; Hunter et al., 2007; Martin et al., 2007; Ellis and Brysbaert, 2010). These findings suggest a predilection of the bilateral hemispheres for segmenting foveal words and projecting them to their contralateral counterparts, as opposed to a simultaneous bilateral projection. This segmentation and subsequent contralateral projection, as posited by SFT, appear to confer computational efficiency, obviating the need for redundant processing across the hemispheres.

Boles (1990) examined a phenomenon of interhemispheric interruption when identical visual stimuli were presented in both the left and right visual fields. This observation intimates that an inhibitory mechanism operates between contralateral hemispheres during the visual recognition of bilaterally presented words, thereby casting doubt on the tenets of the BPT in the context of foveal word recognition. Given the brain's proclivity for computational efficiency, such interhemispheric inhibition can be construed as an adaptive facet of hemispheric regulation. This adaptive mechanism serves to integrate, coordinate, and selectively curate outputs from each hemisphere, a process that is ostensibly essential for the harmonization of the bilateral neural system. Moreover, the empirical inclination toward the SFT in foveal word recognition may be predicated on the avoidance of computational redundancy across the hemispheres, engendered by duplicated projections. Such redundancy not only signifies inefficiency in hemispheric processing but also squanders valuable cognitive resources. Consequently, in the realm of foveal word recognition, the bilateral hemispheres appear to avoid interhemispheric inhibition, likely as a resource-conservation strategy, thereby aligning with the SFT framework wherein words are discretely segmented and projected to their respective contralateral hemispheres.

The present investigation employed a visual half-field presentation paradigm involving both split and identical word presentations to scrutinize the extent to which foveal word recognition aligns with the SFT as opposed to the BPT, particularly in the context of interhemispheric inhibitory regulation. In accordance with the visual half-field presentation paradigm, it is assumed that stimuli presented in the parafoveal region are initially processed by the contralateral hemisphere (Kim et al., 2020, 2022a,b, 2023; Kim and Nam, 2023a,b). Specifically, stimuli appearing in the right visual field (RVF) are initially processed by the left hemisphere (LH), and conversely, stimuli in the left visual field (LVF) are processed by the right hemisphere (RH). We posited the hypothesis that a split presentation of words in the left and right parafoveal visual fields would yield superior performance compared to the simultaneous presentation of identical words in those same fields. This split presentation is postulated to mirror the hemispheric division of labor, thereby aligning with the operational principles of the SFT. Intriguingly, the Korean language serves as an optimal linguistic medium for this line of inquiry, given its rigid syllabic boundaries characterized by either Consonant-Vowel-Consonant (CVC) or Consonant-Vowel (CV) structures, in contrast to the more fluid syllabic configurations found in many Western languages, including English. For instance, the Korean word "책상 (/chek-sang/)" is bifurcated into two distinct syllables, "책 (/chek/)" and "상 (/sang/)," in accordance with Korean's stringent syllabic boundary rules. These syllables are then presented in contralateral visual fields, typically adhering to the left-to-right reading direction ["책 (/chek/)" in the LVF and "상 (/sang/)" in the RVF]. Consequently, the study also incorporated the separate presentation of word syllables to facilitate a comparative analysis with the simultaneous presentation of identical words across bilateral visual fields.

To rigorously interrogate the research hypothesis positing that foveal word visual processing aligns more closely with the SFT than with the BPT—primarily to circumvent interhemispheric inhibition due to redundant processing in the case of identical word projection to both hemispheres—two experiments were executed. The first experiment sought to ascertain whether the visual recognition of Korean words in the present study would also manifest the right visual field advantage (RVFA) and bilateral gain (BG) in laterally presented word recognition, consistent with extant literature. Previous investigations employing lateralized lexical decision tasks with Indo-European languages, notably English, have consistently reported RVFA, indicating superior recognition of words presented in the RVF as opposed to the LVF (Young et al., 1980; Bradshaw and Nettleton, 1983; Hellige, 1993; Mohr et al., 2007). Additionally, BG—defined as enhanced performance in bilaterally presented words relative to unilaterally presented words—has been consistently observed (Mohr et al., 2007). Experiment 1 corroborated the presence of both RVFA and BG in the context of a lateralized lexical decision task using Korean visual words, thereby establishing the suitability of Korean words for visual half-field studies.

In addition, Experiment 2 further delved into the comparative performance between split and identical word recognition in bilateral visual fields, utilizing Korean visual words as the experimental stimuli. We hypothesized that participants would manifest superior performance in split-word presentations relative to identical-word presentations within the bilateral visual field (BVF), a phenomenon attributed to hemispheric inhibitory regulation. To enable this comparative scrutiny, Experiment 2 utilized Korean visual words and assessed performance contrasting split and identical word presentations in the BVF. Furthermore, predicated on the split-fovea theory, we expected superior performance in the responses of split BVF presentations compared to those in the identical presentations in the BVF, and specifically to central visual field (CVF) if there is no corrupted effect from visual acuity, supporting the regulatory interaction between the two hemispheres due to avoid duplication of identical visual stimuli processing. In addition, if this regulatory interhemispheric interaction occurs before word representation stored in mental lexicon, then the benefits in split BVF presentation is observed in both words and pseudowords, meaning interhemispheric regulation in the early stage of visual word processing such as visual-perceptual processing stage. Otherwise, the benefits in split BVF presentation will be shown for words rather than for pseudowords, meaning emergence of regulatory interaction between the two hemispheres in the later stage of visual word processing after lexical access to words in mental lexicon. On the other hand, if the foveal word processing follows processing based on BPT, then we expected superior performance in responses of BVF presentations compared to those in split BVF presentations, meaning advantage from duplicated processing in the two hemisphere. And, likewise, if the benefits in BVF presentations occurs before word presentation, then it would show in both words and pseudowords, meaning the advantage from duplication in both hemispheres occurs irrespective of lexical access to mental lexicon. Otherwise, it would show only in words, meaning the advantage from duplication in both hemispheres only occurs when the stimuli are able to be accessed into mental lexicon.

2 Experiment 1

The primary objective of Experiment 1 was to assess the suitability of Korean visual words within a visual half-field presentation paradigm, focusing on the RVFA and BG. Initially, we posited that visual recognition would be compromised in parafoveal vision relative to foveal vision—a phenomenon termed the 'visual acuity effect'—attributable to the increased viewing angle in parafoveal vision. Given that stimulus clarity generally diminishes with increasing distance from the point of fixation, we anticipated a decline in visual recognition irrespective of the lexicality of the stimulus. Furthermore, we hypothesized that if Korean words are indeed compatible with the visual half-field paradigm, they should exhibit a significant RVFA in parafoveal lexical decision, showing faster and/or more accurate responses for RVF presentation than LVF presentation in words in contrast with in pseudoword. This expectation is grounded in the notion that left-hemispheric dominance in language processing manifests as RVFA in lexical decisions for words as opposed to pseudowords (Knecht et al., 2000; Banich, 2003; Bourne, 2006). In addition, BVF words showed faster and/or more accurate responses for BVF presentation than for RVF presentation in words in contrast with in pseudowords, meaning significant BG only for words. This expectation is grounded in the notion that the co-activation of the bilateral hemisphere in cortical processing by simultaneous parafoveal presentation using identical words is evidenced by BG in lexical decisions for words relative to pseudowords (Hebb, 1949; Mohr et al., 2007).

2.1 Method

2.1.1 Participants

In Experiment 1, a total of 25 participants, all native speakers of Korean, were recruited. The final dataset included all participants, as each adhered to the experimental protocol without exception, yielding a dataset devoid of missing responses or outliers. However, one participant, who registered a score of less than zero on the Edinburgh Handedness Inventory (Oldfield, 1971), was excluded from the final analysis to control for hemispheric asymmetry in language processing based on handedness. The final analytic sample consisted of 13 males and 11 females, with an age distribution of 23.96 ± 2.66 years (M ± SD). Handedness was rigorously controlled, as evidenced by scores on the Edinburgh Handedness Inventory (8.54 ± 1.35). All participants were confirmed to have no visual impairments in either eye and no documented history of mental or physical disabilities. Ethical clearance for the study was obtained from the Institutional Review Board of Korea University, and the study was conducted in strict adherence to the ethical guidelines outlined in the 1964 Declaration of Helsinki. Informed consent was obtained from all participants after they were fully briefed on the study's ethical considerations.

2.1.2 Experimental task

In Experiment 1, participants engaged in a lateralized lexical decision task, wherein they were tasked with discerning whether presented visual letter strings constituted legitimate words or pseudowords. The pseudowords, while orthographically and phonologically valid, lacked semantic content. Stimuli were displayed in one of four visual fields: central (CVF), left (LVF), right (RVF), or both (BVF). The sequence of stimulus presentation was randomized, and participants registered their responses via keyboard input, specifically employing the slash ('/') key for words and the 'z' key for pseudowords. Responses were executed using the index finger of either the left or right hand, with the responding hand counterbalanced across participants. The overarching directive for participants was to render their judgments with both alacrity and precision.

2.1.3 Experiment procedure

The experimental protocol commenced with the display of a fixation point centrally positioned on the screen for a duration of 2000 ms. Upon its disappearance, visual letter strings were presented in one of the designated visual fields—central, left, right, or bilateral—for a temporal window of 180 ms. Participants were then allotted a 2000 ms timeframe within which to categorize the visual letter strings as either words or pseudowords. Prior to embarking on the 400 main trials, which comprised an equal distribution of 200 words and 200 pseudowords, all participants completed 16 practice trials to familiarize themselves with the task. To obviate the potential for stimulus overlap across different visual fields, each stimulus was presented only once throughout the entire experiment, facilitated by the implementation of a Latin square design. In total, four distinct stimulus lists, each containing 200 words and 200 pseudowords, were generated via the Latin square design, with each participant being assigned to one such list.

2.1.4 Apparatus

The experimental stimuli were displayed using an RGB-colored LG monitor situated within a controlled experimental chamber. To ensure a consistent viewing distance, participants were instructed to position their chins on a chin rest, thereby maintaining a fixed 65 cm distance between their nasion and the screen. Furthermore, the visual angles for stimulus presentation were carefully calibrated to fall within a horizontal range of 2° to 5° and a vertical range of 1.5°, in accordance with established guidelines (Ellis et al., 1988; Metusalem et al., 2016). Experimental parameters and stimulus delivery were managed using E-Prime 2.0 professional software (Psychology Software Tools, Inc., Pittsburgh, PA). Participant responses were captured via a keyboard strategically positioned in front of them for ease of data collection.

2.1.5 Materials

In the current experiment, a total of 200 noun words and 200 pseudowords served as the experimental stimuli. For methodological consistency, only two-syllable words and pseudowords were incorporated into the stimulus set. The word stimuli were extracted from the Korean Sejong Corpus, specifically selecting words with a frequency threshold of 10 or higher. Conversely, the pseudowords were constructed by amalgamating syllables present in actual words but were deliberately configured to be undefined within the Korean Sejong Corpus. As a result, these pseudowords were both orthographically and phonologically valid, yet devoid of semantic content.

2.1.6 Experimental conditions

In the experimental design, two primary conditions were manipulated: the visual field of stimulus presentation and lexicality. The visual field condition encompassed presentations in the CVF, RVF, LVF, and BVF, thereby enabling a comparative analysis of response patterns contingent upon the specific visual field in which stimuli were displayed. Lexicality, on the other hand, served as an experimental variable designed to investigate differential responses between legitimate words and pseudowords.

2.1.7 Statistical analyses

In Experiment 1, we performed mixed-effects regression analyses utilizing R software to scrutinize (1) the impact of visual acuity on both RTs and ACC for words and pseudowords, (2) the RVFA on RTs and ACC in words and pseudowords, and (3) the BG on both RTs and ACC for words and pseudowords (R Core Team, 2012). Each analytical model was formulated to incorporate both fixed and random effects, thereby offering a holistic framework for empirical inquiry. Fixed effects encompassed variables of visual field (VF), lexicality, and their two-way interaction (VF × lexicality). The VF delineated into CVF and BVF for examination of the impact of visual acuity, LVF and RVF for examination of the RVFA, and RVF and BVF for examination of the BG. The lexicality delineated into categories of word and pseudoword. Random effects were integrated into the model to account for inter-participant and inter-item variability, thereby ensuring a methodologically rigorous and nuanced analysis. We reported standardized beta values (β), standard errors (*SE*), t statistic, and value of p in the mixed effects regressions for RTs and ACC. The mixed-effect regression models were executed in the R software utilizing the lmer function for RTs and glmer function for ACC.

2.2 Results

Data were acquired for both response times (RTs) and accuracy (ACC) in the context of the lateralized lexical decision task. Preliminary analysis indicated that the ACC for both words and pseudowords across all participants fell within a range of three standard deviations, thus warranting the inclusion of all participant data in the final analysis. The outcomes pertaining to RTs and ACC are delineated in Table 1 and Figure 1, respectively.

2.2.1 Investigation of visual acuity effect in the parafoveal lexical decision using BVF vs. CVF presentation

Initially, the outcomes for RTs revealed significant main effects for both VF [$\beta = -0.119$, $SE = 0.009$, $t = -12.657$, $p < 0.001$] and lexicality [$\beta = 0.243$, $SE = 0.015$, $t = 16.338$, $p < 0.001$]. However, the two-way interaction between VF and lexicality did not attain statistical significance [$\beta = 0.006$, $SE = 0.009$, $t = 0.646$, $p = 0.518$]. The significant main effect of VF suggested accelerated responses in the CVF compared to the BVF. Subsequent analyses of the VF main effect for both words and pseudowords revealed that the significant main effect of VF was attributable to both words [$\beta = -0.141$, $SE = 0.014$, $t = -10.370$, $p < 0.001$] and pseudowords [$\beta = -0.109$, $SE = 0.013$, $t = -8.158$, $p < 0.001$]. Moreover, the significant main effect of lexicality indicated expedited responses for words relative to pseudowords.

Subsequent to the RT analyses, the findings for ACC revealed significant main effects for both VF [$\beta = 0.144$, $SE = 0.030$, $z = 4.786$, $p < 0.001$] and lexicality [$\beta = -0.216$, $SE = 0.057$, $z = -3.816$, $p < 0.001$]. Nonetheless, the two-way interaction between VF and lexicality failed to reach statistical significance [$\beta = -0.028$, $SE = 0.030$, $z = -0.931$, $p = 0.352$]. The pronounced main effect of VF suggested enhanced ACC in the CVF as opposed to the BVF. Further dissection of the VF main effect for both lexical categories—words and pseudowords—indicated that the significant main effect of VF was attributable to both words [$\beta = 0.178$, $SE = 0.046$, $z = 3.879$, $p < 0.001$] and pseudowords [$\beta = 0.120$, $SE = 0.040$, $z = 2.982$, $p = 0.003$]. Moreover, the significant main effect of lexicality denoted superior ACC for words in comparison to pseudowords.

2.2.2 Investigation of RVFA in the parafoveal lexical decision using LVF vs. RVF presentation

An initial analysis focused on RTs and the analysis yielded a significant main effect for lexicality [$\beta = 0.244$, $SE = 0.015$, $t = 16.447$, $p < 0.001$], as well as a noteworthy two-way interaction between VF and lexicality [$\beta = 0.021$, $SE = 0.010$, $t = 2.231$, $p = 0.026$]. In contrast, the main effect associated with VF did not attain statistical significance [$\beta = -0.001$, $SE = 0.010$, $t = -0.100$, $p = 0.920$]. The pronounced main effect for lexicality suggested more rapid RTs for words as compared to pseudowords. Subsequent exploration of the significant interaction between VF and lexicality through simple main effect analysis revealed that neither the effect of VF for words [$\beta = -0.026$, $SE = 0.014$, $t = -1.902$, $p = 0.057$] nor for pseudowords [$\beta = 0.019$, $SE = 0.013$, $t = 1.402$, $p = 0.161$] reached statistical significance.

Subsequent to the evaluation of RTs, the analysis was extended to examine ACC. The statistical output revealed significant main effects

TABLE 1 Results of the response times (RT) and accuracy rates (ACC) in the lateralized lexical decision task in Experiment 1.

	CVF		BVF		RVF		LVF	
	RT	ACC	RT	ACC	RT	AC	RT	AC
Words	600 (74)	0.907 (0.077)	662 (72)	0.843(0.099)	684 (73)	0.839(0.102)	679 (98)	0.823(0.102)
Pseudowords	632 (116)	0.886(0.098)	686 (128)	0.856(0.122)	704 (128)	0.786(0.119)	713 (115)	0.759(0.162)

The bracket value indicates the standard deviation.

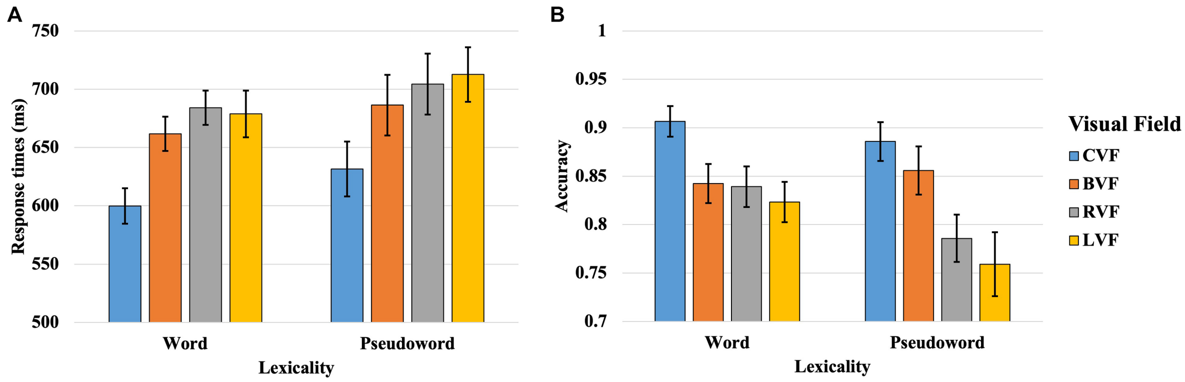


FIGURE 1
Results of the response times (**A**) and accuracy (**B**) in the CVF, BVF, RVF, and LVF in lateralized lexical decision task of Experiment 1. The line in the bar indicates standard error. The standard error is computed by dividing the standard deviation by the square root of the sample size. The standard error serves to assess how closely a statistic derived from the sample approximates the true parameters of the overall population.

for both VF [$\beta = 0.071$, $SE = 0.030$, $z = 2.363$, $p = 0.018$] and lexicality [$\beta = -0.215$, $SE = 0.056$, $z = -3.814$, $p < 0.001$]. However, the interaction between VF and lexicality did not reach statistical significance [$\beta = -0.051$, $SE = 0.030$, $z = -1.708$, $p = 0.088$]. The main effect for VF suggested a heightened level of ACC in the RVF compared to the LVF. Upon disaggregating the VF effect by word and pseudoword categories, it was observed that the VF effect was primarily driven by words [$\beta = 0.125$, $SE = 0.046$, $z = 2.735$, $p = 0.006$], rather than pseudowords [$\beta = 0.021$, $SE = 0.040$, $z = 0.527$, $p = 0.598$]. Additionally, the main effect of lexicality indicated superior ACC for words relative to pseudowords.

2.2.3 Investigation of BG in the parafoveal lexical decision using BVF vs. RVF presentation

Initial analyses were executed on RTs. The outcomes revealed salient main effects for both VF [$\beta = 0.032$, $SE = 0.010$, $t = 3.355$, $p < 0.001$] and lexicality [$\beta = 0.243$, $SE = 0.015$, $t = 16.431$, $p < 0.001$]. Contrarily, the two-way interaction between VF and lexicality did not attain statistical significance [$\beta = -0.016$, $SE = 0.010$, $t = -1.721$, $p = 0.085$]. The pronounced main effect for VF suggested expedited responses in the BVF compared to the RVF. Subsequent analyses partitioning the VF main effect by word and pseudoword categories revealed that the significant VF effect was attributable to words [$\beta = 0.053$, $SE = 0.014$, $t = 3.841$, $p < 0.001$], rather than pseudowords [$\beta = 0.017$, $SE = 0.013$, $t = 1.280$, $p = 0.201$]. Additionally, the main effect of lexicality indicated accelerated responses for words relative to pseudowords.

Subsequent to the RTs analysis, the findings for ACC revealed robust main effects for both VF [$\beta = -0.120$, $SE = 0.030$, $z = -3.994$, $p < 0.001$] and lexicality [$\beta = -0.218$, $SE = 0.057$, $z = -3.849$, $p < 0.001$]. In contrast, the two-way interaction between VF and lexicality did not reach statistical significance [$\beta = 0.057$, $SE = 0.030$, $z = 1.882$, $p = 0.060$]. The pronounced main effect for VF suggested enhanced ACC in the BVF as opposed to the RVF. Further dissection of the VF main effect by word and pseudoword categories indicated that the significant VF effect was attributable to words [$\beta = -0.192$, $SE = 0.046$, $z = -4.168$, $p < 0.001$], but not to pseudowords [$\beta = -0.063$, $SE = 0.040$, $z = -1.561$, $p = 0.118$]. Additionally, the main effect of lexicality underscored superior ACC for words relative to pseudowords.

2.3 Discussion

Experiment 1 aimed to examine the RVFA and BG in a lateralized lexical decision paradigm utilizing Korean visual words, in alignment with existing scholarly contributions (e.g., Mohr et al., 2007). Initially, the study revealed a pronounced visual acuity effect for both words and pseudowords, characterized by attenuated speed and ACC for parafoveal stimuli compared to foveal stimuli, irrespective of lexicality. This observation substantiates the notion of a decremental effect in parafoveal lexical decision-making, attributable to the increased viewing angle, thereby validating the visual half-field experimental design. Furthermore, the data corroborated significant RVFA and BG phenomena in the context of Korean visual word recognition, thereby replicating previous findings in other languages such as the RVFA in English (e.g., Barca et al., 2011), the BG in German (e.g., Mohr et al., 2007). The manifestation of RVFA implies a left-hemispheric predominance in the processing of Korean visual words (Knecht et al., 2000; Banich, 2003; Bourne, 2006), while the presence of BG suggests interhemispheric facilitation during bilateral word presentation (Mohr et al., 2007), in contrast to pseudoword conditions.

The presence of RVFA and BG in Korean, a language characterized by multisyllabic words, intimates that these phenomena are not contingent upon the morphological attributes of the words. This observation suggests the potential generalization of RVFA and BG in parafoveal word recognition across diverse linguistic architectures, including agglutinative (e.g., Korean) and alphabetic (e.g., English) languages. The consistency of RVFA and BG effects across languages suggests that language-specific traits, such as morphological structure, do not exert a significant influence on parafoveal word recognition. This universality underscores the left-hemispheric dominance and bilateral hemispheric cooperation in language processing, thereby affirming the methodological aptness of employing Korean visual words, particularly in parafoveal presentations, for future inquiries into hemispheric division of labor.

Furthermore, the strict syllabic demarcation inherent to Korean words offers a unique opportunity for subsequent experiments. Specifically, in Experiment 2, the use of Korean words will facilitate the exploration of interhemispheric inhibition through the manipulation of split-word presentations and the simultaneous display of identical words in the bilateral visual field. This is particularly pertinent for investigating interhemispheric inhibition predicated on the split-fovea theory, a manipulation that is more challenging to implement in languages like English, where syllabic boundaries are less rigidly defined.

3 Experiment 2

Experiment 2 aimed to investigate whether the foveal word recognition follows SFT rather than BPT due to a mechanism to mitigate interhemispheric inhibition arising from redundant processing during identical word presentations to contralateral hemispheres. We hypothesized that participants would manifest superior performance in split-word presentations relative to identical-word presentations within the BVF, a phenomenon attributed to hemispheric inhibitory regulation. To enable this comparative scrutiny, the experiment utilized Korean visual words and assessed performance contrasting split and identical word presentations in the BVF. Furthermore, predicated on the split-fovea theory, the study sought to compare the response of split BVF presentations with those in the CVF. This comparison was designed to discern whether the disparities between split and identical BVF presentation would endure when contrasting parafoveal split BVF processing with foveal central word processing. Should visual acuity effects persist in diminishing performance in split BVF processing, a notable divergence between split BVF and CVF lexical decisions is anticipated. Conversely, if split BVF processing aligns with the assumptions of SFT, irrespective of any decremental visual acuity effects, no significant difference between split BVF and CVF outcomes is expected.

3.1 Method

3.1.1 Participants

In Experiment 2, an initial cohort of 43 native Korean speakers was recruited. One participant was subsequently excluded from the final data analysis due to non-compliance with experimental procedures, resulting in a final sample of 42 participants (15 males and 27 females; age 25.21 ± 4.03 years, M ± SD). Handedness was controlled across the sample, as evidenced by scores on the Edinburgh Handedness Inventory (8.19 ± 1.93) (Oldfield, 1971). All participants reported no visual impairments in either eye and had no documented history of mental or physical disabilities. Ethical approval for the study was granted by the Institutional Review Board of Korea University, Korea, where the research was conducted. The study was executed in strict compliance with the ethical guidelines stipulated in the 1964 Declaration of Helsinki. All participants were apprised of the ethical considerations and provided informed consent prior to their involvement.

3.1.2 Experimental task

Experiment 2 also employed a lateralized lexical decision task, wherein participants were tasked with categorizing visual strings as either legitimate words or pseudowords. Notably, the pseudowords in this experiment were both orthographically valid and pronounceable, yet imbued with semantic content. Stimuli were displayed either in the CVF or the BVF. Within the BVF condition, two distinct types of presentations were utilized. The first entailed a simultaneous presentation of identical stimuli in the BVF; for instance, participants were exposed to the identical word '학교 (/hak-kyo/)' in both the left and right visual fields concurrently. Conversely, the second type involved a split presentation in the BVF, wherein the word '학교 (/hak-kyo/)' was bifurcated into its constituent syllables '학 (/hak/)' and '교 (/kyo/)', each of which was displayed separately in either the left or right visual field. The sequence of stimulus presentation was randomized, and participants registered their judgments via keyboard input, specifically employing the slash ('/') key for words and the 'z' key for pseudowords. Responses were executed using the index finger of either the left or right hand, with the responding hand counterbalanced across participants. The overarching directive for participants was to render their judgments with both alacrity and precision.

3.1.3 Experimental procedure

The experimental protocol for Experiment 2 commenced with a centrally positioned fixation point displayed on the screen for a duration of 2000 ms. Subsequent to this, visual letter strings were presented either in the CVF or in one of two types of BVFs for a temporal window of 180 ms. Participants were allotted a 2000 ms timeframe within which to categorize these visual letter strings as either words or pseudowords. The complete procedural outline of Experiment 2 is delineated in Figure 2. Prior to embarking on the 396 main trials, which comprised an equal distribution of 198 words and 198 pseudowords, participants completed 12 practice trials for task familiarization. To mitigate the risk of stimulus overlap across different visual fields, each stimulus was presented only once throughout the experiment, facilitated by the implementation of a Latin square design. Consequently, three distinct stimulus lists, each containing 198 words and 198 pseudowords, were generated via the Latin square design, with each participant being assigned to one such list.

3.1.4 Apparatus

Consistent with the methodology employed in Experiment 1, participants were subjected to the same experimental protocol.

3.1.5 Materials

In Experiment 2, the stimulus set was derived from the materials utilized in Experiment 1, with the exclusion of two noun words and two pseudowords to align with the experimental design of the current study.

3.1.6 Experimental conditions

In the experimental framework of Experiment 2, two primary conditions were manipulated: the visual field of stimulus presentation and lexicality. The visual field condition encompassed presentations in the CVF, as well as two types of presentations in the BVF—simultaneous and split. These variations facilitated a nuanced comparison of response patterns contingent upon the specific visual field in which stimuli were displayed. Lexicality served as an additional experimental variable, designed to examine differential responses between legitimate words and pseudowords.

3.1.7 Statistical analyses

In Experiment 2, we performed mixed-effects regression analyses utilizing R software to scrutinize (1) the differential

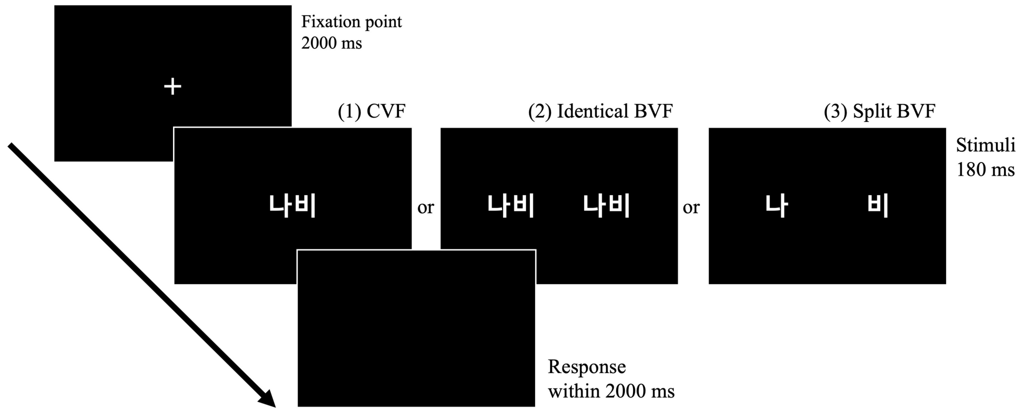


FIGURE 2
Schematic of the experimental paradigm employed in Experiment 2 for lateralized lexical decision task, illustrating stimulus presentation modalities in the CVF, identical bilateral visual field (identical BVF), and split bilateral visual field (split BVF).

impact of split and identical BVFs on RTs and ACC in the context of words and pseudowords, and (2) the differential impact of split BVF and CVF on RTs and ACC across words and pseudowords (R Core Team, 2012). Each analytical model was formulated to incorporate both fixed and random effects, thereby offering a holistic framework for empirical inquiry. Fixed effects encompassed variables of visual field (VF), lexicality, and their two-way interaction (VF × lexicality). The VF delineated into split BVF and identical BVF for examination of the differential impact of split and identical BVFs, split BVF and CVF for examination of the differential impact of split BVF and CVF. The lexicality delineated into categories of word and pseudoword. Random effects were integrated into the model to account for inter-participant and inter-item variability, thereby ensuring a methodologically rigorous and nuanced analysis. The mixed-effect regression models were executed in the R software utilizing the lmer function for RTs and glmer function for ACC.

3.2 Results

Data were amassed for both RTs and ACC in the context of a lateralized lexical decision task. Preliminary preprocessing analysis indicated that ACC metrics for both words and pseudowords were confined within a three-standard-deviation range for the entire participant pool, save for two outliers. To maintain the analytical robustness and integrity of the study, these two exceptional datasets were omitted from the final evaluation. The synthesized outcomes, delineated in Table 2 and Figure 3, expound upon the RT and ACC parameters observed in Experiment 2.

3.2.1 Investigation of SFT in lexical decision using split vs. identical BVF presentation

An initial analysis targeting RTs found statistically significant main effects for both VF [$\beta = -0.022$, $SE = 0.007$, $t = -3.326$, $p < 0.001$] and lexicality [$\beta = 0.235$, $SE = 0.015$, $t = 15.828$, $p < 0.001$]. Conversely, the interaction between VF and lexicality failed to reach statistical significance [$\beta = -0.002$, $SE = 0.007$, $t = -0.371$, $p = 0.711$]. The main effect of VF suggested accelerated RTs in the split BVF condition compared to the simultaneous BVF condition. Subsequent analysis of the VF main effect revealed that this acceleration was observed both for words [$\beta = -0.021$, $SE = 0.009$, $t = -2.271$, $p = 0.023$] and pseudowords [$\beta = -0.024$, $SE = 0.010$, $t = -2.535$, $p = 0.011$], indicating that the split BVF condition facilitated faster responses irrespective of stimulus lexicality. Additionally, the main effect of lexicality indicated a response time advantage for words over pseudowords.

Subsequent to the RT analysis, the analysis for ACC revealed statistically significant main effects for both VF [$\beta = 0.066$, $SE = 0.029$, $z = 2.233$, $p = 0.026$] and lexicality [$\beta = -0.380$, $SE = 0.069$, $z = -5.513$, $p < 0.001$]. Notably, a significant two-way interaction between VF and lexicality was also observed [$\beta = 0.173$, $SE = 0.029$, $z = 5.862$, $p < 0.001$]. The main effect of VF suggested enhanced ACC in the split BVF condition relative to the identical BVF condition. Concurrently, the main effect of lexicality indicated superior ACC for pseudowords compared to words. Further dissection of the significant VF × lexicality interaction revealed a significant simple main effect of VF for words [$\beta = -0.130$, $SE = 0.047$, $z = -2.734$, $p = 0.006$], signifying greater ACC in the split BVF condition. Likewise, a significant simple main effect of VF was found for pseudowords [$\beta = 0.245$, $SE = 0.037$, $z = 6.680$, $p < 0.001$], also indicating enhanced ACC in the split BVF condition.

3.2.2 Investigation of visual acuity effect in the parafoveal lexical decision using split BVF vs. CVF presentation

An initial analysis of RTs yielded significant main effects for both VF [$\beta = 0.136$, $SE = 0.007$, $t = 20.816$, $p < 0.001$] and lexicality [$\beta = 0.233$, $SE = 0.015$, $t = 15.616$, $p < 0.001$]. Furthermore, a significant two-way interaction between VF and lexicality was observed [$\beta = -0.018$, $SE = 0.007$, $t = -2.821$, $p = 0.005$]. The main effect of VF revealed expedited responses in the CVF compared to the split BVF. Additionally, the main effect of lexicality indicated accelerated responses for words relative to pseudowords. Subsequent simple main effect analyses on the significant VF × lexicality interaction disclosed that the simple main effect of VF was significant for both words [$\beta = 0.171$, $SE = 0.009$, $t = 18.657$, $p < 0.001$] and pseudowords [$\beta = 0.113$, $SE = 0.010$, $t = 11.910$, $p < 0.001$], signifying more rapid responses in the CVF compared to the split BVF, irrespective of lexicality.

TABLE 2 Results of the response times (RT) and accuracy rates (ACC) in the lateralized lexical decision task in Experiment 2.

	CVF		Split BVF		Identical BVF	
	RT	ACC	RT	ACC	RT	ACC
Words	614 (96)	0.929(0.057)	675 (106)	0.891(0.063)	688 (121)	0.857(0.104)
Pseudowords	619 (97)	0.924(0.063)	681 (103)	0.898(0.064)	688 (102)	0.891(0.070)

The bracket value indicates the standard deviation.

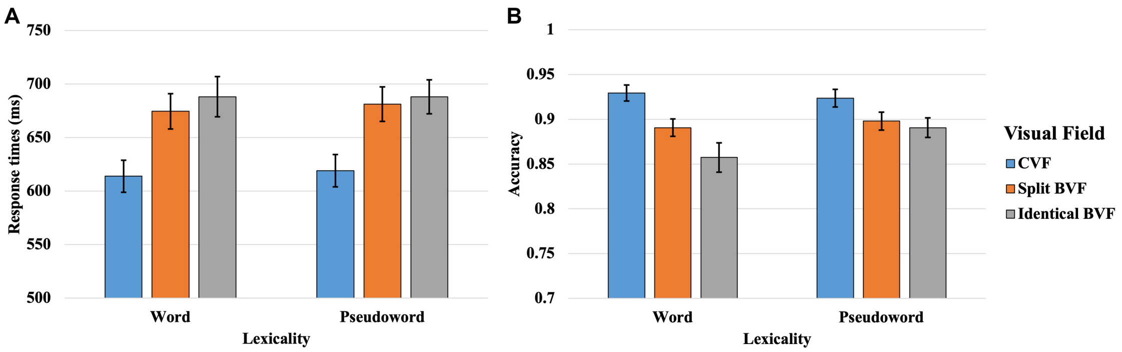


FIGURE 3
Results of the response times (**A**) and accuracy (**B**) in CVF, split BVF, and identical BVF in lateralized lexical decision task of Experiment 2. The line in the bar indicates standard error.

Subsequent to the RT analysis, the statistical outcomes for ACC revealed salient main effects for both VF [$\beta = -0.238$, $SE = 0.031$, $z = -7.761$, $p < 0.001$] and lexicality [$\beta = -0.433$, $SE = 0.070$, $z = -6.210$, $p < 0.001$]. Additionally, a significant two-way interaction between VF and lexicality was observed [$\beta = 0.238$, $SE = 0.031$, $z = 7.775$, $p < 0.001$]. The main effect of VF demonstrated enhanced ACC in the CVF as compared to the split BVF. Concurrently, the main effect of lexicality indicated superior ACC for pseudowords relative to words. A subsequent simple main effect analysis on the VF × lexicality interaction disclosed a significant simple main effect of VF for words [$\beta = -0.498$, $SE = 0.052$, $z = -9.658$, $p < 0.001$], signifying heightened ACC in the CVF over the split BVF. However, the simple main effect of VF for pseudowords was not statistically significant [$\beta = 0.002$, $SE = 0.036$, $z = 0.065$, $p = 0.948$], signifying no difference of ACC between the CVF and the split BVF presentations.

3.3 Discussion

Experiment 2 revealed that RTs were slower and ACC was diminished in the identical BVF as compared to the split BVF, irrespective of target lexicality. These findings suggest a superior visual recognition performance in the split BVF, lending empirical support to the split-fovea theory. Additionally, a significant performance discrepancy was observed between the split BVF and the CVF. This indicates that, despite the benefits of split presentation at parafoveal vision, a degradation in performance persists, attributable to the limitations of visual acuity in parafoveal presentations.

The findings of Experiment 2 corroborate extant literature, such as the work of Chiarello and Maxfield (1996), which posits the occurrence of interhemispheric inhibition during simultaneous presentation of identical words in both the left and right visual fields. This is evidenced by the slower RTs observed in the BVF in our study. Such inhibitory regulation between the hemispheres is postulated to serve as a compensatory mechanism aimed at mitigating redundant processing across both hemispheres. Given the brain's proclivity for efficiency in cognitive processing, particularly in the context of mental energy conservation, interhemispheric inhibition serves to judiciously allocate limited neural resources. This shows the superior recognition performance for split words as compared to the simultaneous presentation of identical words in the BVFs.

An additional intriguing outcome of Experiment 2 was the absence of a significant delay in RTs for pseudoword processing in the BVF as compared to the split BVF, particularly when contrasted with word processing. This lack of delay in pseudoword processing suggests that interhemispheric inhibition in visual recognition is contingent upon lexical access to the mental lexicon, which is rather later stage of visual word processing such as lexical processing after visual-perceptual processing. This phenomenon can be attributed to hemispheric competition that arises during lexical access in the context of identical word recognition in the BVF. Such competition is engendered by the shared pathway for accessing the mental lexicon from both the left and right hemispheres.

The findings of this study lend empirical support to the SFT over the BPT in the context of foveal word recognition. The observed hemispheric conflicts, engendered by interhemispheric inhibition in BVF word recognition, suggest a predilection for SFT-based processing over BPT in foveal word recognition. When identical words are projected in the BVF, each contralateral hemisphere is activated to process the words via a shared lexical access pathway to the mental lexicon. This activation engenders hemispheric conflicts during lexical access, likely as a metabolic conservation strategy to mitigate the redundancy inherent in simultaneous activation of both hemispheres. In contrast, such conflicts are conspicuously absent in split BVF word recognition. In this condition, each hemisphere processes a distinct syllabic component of the word in the unilateral visual fields (UVFs), which are subsequently integrated to form a complete word. This obviates the need for redundant processing and the associated metabolic costs, thereby eliminating the delays observed in BVF word recognition. Thus, the superior performance in split BVF word recognition relative to BVF word recognition can be attributed to the mitigation of hemispheric conflict through interhemispheric inhibition, reinforcing the primacy of SFT in foveal word recognition.

4 General discussion

The present investigation conducted two experiments to ascertain whether foveal word recognition adheres more closely to the SFT than to the BPT, with a focus on the role of interhemispheric inhibitory regulation. Experiment 1 demonstrated the presence of RVFA and BG in a lateralized lexical decision task using Korean words. This outcome substantiates the feasibility of employing Korean words in visual half-field studies, akin to research conducted in other languages such as English. Experiment 2 revealed accelerated RTs in the visual recognition of split words presented in the BVF as compared to identical words also presented in the BVF. This finding suggests that the bilateral hemispheres engage in a division of labor to circumvent interhemispheric inhibition, particularly when identical words are propagated to both hemispheres, as opposed to split words in the BVF.

Indeed, interhemispheric inhibition serves as a critical regulatory mechanism for dynamic hemispheric processing within the brain. Chiarello and Maxfield (1996) delineated three distinct forms of interhemispheric inhibition. The first form entails functional suppression, wherein one hemisphere exerts inhibitory control over its contralateral counterpart during cognitive processing (e.g., Cook, 1984). Previous research elucidating this suppressive interaction posits that hemispheric dominance is achieved by one hemisphere inhibiting the other, thereby reducing parallel processing and mitigating potential conflicts between the hemispheres. The second form of inhibition is characterized by hemispheric isolation, aimed at alleviating potential bottlenecks in interhemispheric interactions (e.g., Zaidel et al., 1990; Hellige, 1993). This form is distinct from the first in that it allows for parallel processing within each hemisphere. Here, the interhemispheric transfer pathway is inhibited, effectively blocking communication between the hemispheres. While this blockade precludes interhemispheric interactions, it permits each hemisphere to function in parallel, thereby isolating them from each other. The third form of inhibition diverges from both functional suppression and hemispheric isolation, focusing instead on the restriction of one hemisphere's efficiency by the other (Liederman, 1986; Clarke et al., 1993). This manifests as interhemispheric interference, wherein each hemisphere is presented with irrelevant or distracting information via the cortical pathways that facilitate interaction between the two hemispheres.

These three modalities of interhemispheric inhibition are posited to be instrumental in sustaining a harmonious and adaptive neural processing framework. Such inhibitory mechanisms between the hemispheres facilitate optimized hemispheric responses by mitigating redundancy. Given that duplicative processing across the left and right hemispheres is superfluous, it is plausible that one hemisphere exerts regulatory control over its contralateral counterpart to minimize redundant neural activations, particularly in the context of identical word projections to the BVFs.

This regulatory interplay between the left and right hemispheres may manifest as functional differentiation, potentially giving rise to hemispheric specialization—for instance, the left hemisphere's dominance in language processing. Such regulatory mechanisms serve as a framework for the functional partitioning of tasks across the hemispheres. In this context, Karbe et al. (1998) investigated the influence of transcallosal inhibitory activity on functional brain asymmetry, employing three-dimensional magnetic resonance imaging for metabolic assessments. Their findings revealed metabolic alterations in the midbody of the corpus callosum and isthmus, which exhibited a negative correlation with activity in language-associated regions such as the left inferior cortex and the right superior temporal cortex. Specifically, as metabolic activity in the midbody of the corpus callosum increased, metabolic activity in these asymmetrically functioning cortical areas decreased. These observations suggest that interhemispheric inhibition, mediated through callosal fiber tracts, is intricately linked with the functional asymmetries observed between the left and right cortical regions. Such functional disparities serve to reinforce hemispheric lateralization or specialization in specific cognitive tasks, such as language processing predominantly governed by the left hemisphere.

Consequently, regulatory mechanisms between the left and right hemispheres facilitate a division of labor that optimizes foveal word recognition. This hemispheric partitioning enhances processing efficiency by mitigating superfluous interhemispheric inhibition. Such autonomous functioning of each hemisphere serves to preempt potential conflicts between the hemispheres, thereby streamlining cognitive processing.

In summary, the present investigation conducted two experiments to examine the mechanisms of foveal word recognition through the lens of hemispheric interactions. The findings revealed suboptimal performance in split word presentation compared to identical word presentation in the BVFs, implicating a division of labor across the hemispheres. This division appears to be driven by the need to circumvent inhibitory regulation that arises from simultaneous propagation of identical words to both hemispheres. By adhering to this hemispheric specialization, the bilateral processing of foveal words is consequently enhanced.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary materials, further inquiries can be directed to the corresponding author.

Ethics statement

The studies involving humans were approved by Korea University Institutional Review Board (KUIRB-2021-0427-01). The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

SK: Conceptualization, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Visualization, Writing – original draft, Writing – review & editing. KN: Supervision, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This research was supported by the MSIT (Ministry of Science and ICT), South Korea, under the ITRC (Information Technology Research Center) support program (IITP-2022-2017-0-01630) supervised by the IITP (Institute for Information and communications Technology Promotion) and by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2022S1A5B5A16049384).

Conflict of interest

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

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2023.1293529/full#supplementary-material

References

Banich, M. T. (2003). "Interaction between the hemispheres and its implications for the processing capacity of the brain," in *Brain asymmetry* (2nd ed.). eds. R. Davidson and K. Hugdahl (Cambridge, MA: MIT Press), 261–302.

Barca, L., Cornelissen, P., Simpson, M., Urooj, U., Woods, W., and Ellis, A. W. (2011). The neural basis of the right visual field advantage in reading: an MEG analysis using virtual electrodes. *Brain Lang.* 118, 53–71. doi: 10.1016/j.bandl.2010.09.003

Boles, D. B. (1990). What bilateral displays do. *Brain Cogn.* 12, 205–228. doi: 10.1016/0278-2626(90)90016-H

Bourne, V. J. (2006). The divided visual field paradigm: methodological considerations. *Laterality* 11, 373–393. doi: 10.1080/13576500600633982

Bradshaw, J. L., and Nettleton, N. C. (1983). Hemispheric specialization: return to a house divided. *Behav. Brain Sci.* 6, 528–533. doi: 10.1017/S0140525X00073775

Brysbaert, M. (2004). The importance of interhemispheric transfer for foveal vision: a factor that has been overlooked in theories of visual word recognition and object perception. *Brain Lang.* 88, 259–267. doi: 10.1016/S0093-934X(03)00279-7

Brysbaert, M., Vitu, F., and Schroyens, W. (1996). The right visual field advantage and the optimal viewing position effect: on the relation between foveal and parafoveal word recognition. *Neuropsychology* 10:385. doi: 10.1037/0894-4105.10.3.385

Bunt, A. H., Minckler, D. S., and Johanson, G. W. (1977). Demonstration of bilateral projection of the central retina of the monkey with horseradish peroxidase neuronography. *J. Comp. Neurol.* 171, 619–630. doi: 10.1002/cne.901710412

Chiarello, C., and Maxfield, L. (1996). Varieties of interhemispheric inhibition, or how to keep a good hemisphere down. *Brain Cogn.* 30, 81–108. doi: 10.1006/brcg.1996.0006

Clarke, J. M., Lufkin, R. B., and Zaidel, E. (1993). Corpus callosum morphometry and dichotic listening performace: Individual differences in functional interhemispheric inhibition? *Neuropsychologia* 31, 547–557.

Cook, N. D. (1984). Callosal inhibition: the key to the brain code. *Behav. Sci.* 29, 98–110. doi: 10.1002/bs.3830290203

Ellis, A. W., and Brysbaert, M. (2010). Split fovea theory and the role of the two cerebral hemispheres in reading: a review of the evidence. *Neuropsychologia* 48, 353–365. doi: 10.1016/j.neuropsychologia.2009.08.021

Ellis, A. W., Young, A. W., and Anderson, C. (1988). Modes of word recognition in the left and right cerebral hemispheres. *Brain Lang.* 35, 254–273.

Hebb, D. (1949). *The organization of behavior. A neuropsychological theory*. New York: Wiley

Hellige, J. B. (1993). Unity of thought and action: varieties of interaction between the left and right cerebral hemispheres. *Curr. Dir. Psychol. Sci.* 2, 21–26. doi: 10.1111/1467-8721.ep10770559

Hunter, Z. R., Brysbaert, M., and Knecht, S. (2007). Foveal word reading requires interhemispheric communication. *J. Cogn. Neurosci.* 19, 1373–1387. doi: 10.1162/jocn.2007.19.8.1373

Iacoboni, M., and Zaidel, E. (1996). Hemispheric independence in word recognition: evidence from unilateral and bilateral presentations. *Brain Lang.* 53, 121–140. doi: 10.1006/brln.1996.0040

Karbe, H., Herholz, K., Halber, M., and Heiss, W. D. (1998). Collateral inhibition of transcallosal activity facilitates functional brain asymmetry. *J. Cereb. Blood Flow Metab.* 18, 1157–1161.

Kim, S., Kim, J., and Nam, K. (2022a). Familiarity with words modulates interhemispheric interactions in visual word recognition. *Front. Psychol.* 13:892858. doi: 10.3389/fpsyg.2022.892858

Kim, S., Koo, M., Kim, J., and Nam, K. (2020). The research for language information processing of bilateral hemispheres on Korean noun Eojeol: visual half-field study. *Korean J. Cognit. Biol. Psychol.* 32, 29–53. doi: 10.22172/cogbio.2020.32.1.003

Kim, S., Lee, C., and Nam, K. (2022b). The examination of the visual-perceptual locus in hemispheric laterality of the word length effect using Korean visual word. *Laterality* 27, 485–512. doi: 10.1080/1357650X.2022.2103144

Kim, S., and Nam, K. (2023a). Asymmetries in hemispheric strategies for visual recognition of homonyms. *Laterality* 28, 4–6. doi: 10.1080/1357650X.2023

Kim, S., and Nam, K. (2023b). Examining interhemispheric processing and task demand in lexical decision-making: insights from lateralized visual field paradigm. *Front. Psychol.* 14:2317. doi: 10.3389/fpsyg.2023.1208786

Kim, S., Song, J., Lee, W., and Nam, K. (2023). The patterns of intra−/inter-hemispheric interactions of left and right hemispheres in visual word processing. *Cogn. Neurosci.* 14, 137–151. doi: 10.1080/17588928.2023.2259555

Knecht, S., Dräger, B., Deppe, M., Bobe, L., Lohmann, H., Flöel, A., et al. (2000). Handedness and hemispheric language dominance in healthy humans. *Brain* 123, 2512–2518. doi: 10.1093/brain/123.12.2512

Lavidor, M., and Walsh, V. (2004). The nature of foveal representation. *Nat. Rev. Neurosci.* 5, 729–735. doi: 10.1038/nrn1498

Liederman, J. (1986). Interhemispheric interference during word naming. *Int. J. Neurosci.* 30, 43–56.

Lindell, A. K., Arend, I., Ward, R., Norton, J., and Wathan, J. (2007). Hemispheric asymmetries in feature integration during visual word recognition. *Laterality* 12, 543–558. doi: 10.1080/13576500701495190

Martin, C. D., Thierry, G., Démonet, J. F., Roberts, M., and Nazir, T. (2007). ERP evidence for the split fovea theory. *Brain Res.* 1185, 212–220. doi: 10.1016/j.brainres.2007.09.049

Metusalem, R., Kutas, M., Urbach, T. P., and Elman, J. L. (2016). Hemispheric asymmetry in event knowledge activation during incremental language comprehension: A visual half-field ERP study. *Neuropsychologia* 84, 252–271.

Mohr, B., Endrass, T., Hauk, O., and Pulvermüller, F. (2007). ERP correlates of the bilateral redundancy gain for words. *Neuropsychologia* 45, 2114–2124. doi: 10.1016/j.neuropsychologia.2007.01.015

Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* 9, 97–113. doi: 10.1016/0028-3932(71)90067-4

Portin, K., Salenius, S., Salmelin, R., and Hari, R. (1998). Activation of the human occipital and parietal cortex by pattern and luminance stimuli: neuromagnetic measurements. *Cereb. Cortex* 8, 253–260. doi: 10.1093/cercor/8.3.253

R Core Team (2012). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing: Vienna, Austria.

Young, A. W., Bion, P. J., and Ellis, A. W. (1980). Studies toward a model of laterality effects for picture and word naming. *Brain Lang.* 11, 54–65. doi: 10.1016/0093-934X(80)90109-1

Zaidel, E., Clarke, J. M., and Suyenobu, B. (1990). "Hemispheric independence: a paradigm case for cognitive neuroscience" in *Neurobiology of higher cognitive functions*. eds. A. B. Scheibel and A. F. Wechsler (New York: Guilford Press), 297–355.

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 21 February 2024
DOI 10.3389/fpsyg.2024.1281926

OPEN ACCESS

EDITED BY
Katarzyna Jankowiak,
Adam Mickiewicz University, Poland

REVIEWED BY
Yue Yu,
Southeast University, China
Rüdiger Christoph Pryss,
Julius Maximilian University of Würzburg, Germany

*CORRESPONDENCE
Haixiao Wang
hxwang@nju.edu.cn

RECEIVED 23 August 2023
ACCEPTED 08 February 2024
PUBLISHED 21 February 2024

CITATION
Fan K and Wang H (2024) The productive processing of formulaic sequences by second language learners in writing.
Front. Psychol. 15:1281926.
doi: 10.3389/fpsyg.2024.1281926

The productive processing of formulaic sequences by second language learners in writing

Kunmeng Fan[1,2] and Haixiao Wang[3]*

[1]School of Foreign Studies, Nanjing University, Nanjing, China, [2]School of Foreign Languages, Anhui University of Technology, Ma'anshan, China, [3]Department of Applied Foreign Language Studies, Nanjing University, Nanjing, China

There has been much debate in psycholinguistic research on whether formulaic sequences (FSs) are processed holistically or in a compositional manner. Whereas most previous studies on this issue focused on the receptive processing of FSs, few have investigated the productive processing of FSs, particularly in the second language (L2) learning context. Besides, most previous studies on L2 FSs examined learner-external FSs, or those identified by external criteria such as corpus frequency with little attention to learner-internal FSs, or psychological units perceived as wholes by learners themselves, although there might be much overlap between learner-external and learner-internal FSs. This study was designed to explore the productive processing of FSs by L2 learners from their own perspective, while taking into account the effects of L2 proficiency and topic familiarity. It made a distinction between internal FSs and purely external FSs as the primary criterion of categorizing learners' processing behaviors. Ten Chinese English learners from two proficiency levels completed two writing tasks differing in topic familiarity. Upon the completion of each task, each participant and the researcher identified the FSs separately and then distinguished internal FSs and purely external FSs (termed as assembled FSs, since they were perceived as being assembled from scratch) collectively. Next, each participant performed video stimulated recall (VSR) for the production process of each FS. The results showed that the learners' conscious processing (i.e., retrieval/assembly and integration into the text) of FSs can be categorized on two levels (lexical and syntactic). There was more holistic processing than compositional processing on the lexical level, but not on the syntactic level, indicating the learners' sizable storages of FSs and the syntactic flexibility of FSs. Furthermore, between-group differences and between-task differences were detected on two processing levels: higher-proficiency students retrieved more internal FSs and made more modifications to them than their lower-proficiency counterparts; in the familiar-topic writing, learners retrieved more internal FSs and made less modifications to them. Based on the findings, a model of L2 FS production is proposed, and pedagogical implications for the teaching of L2 FSs are provided.

KEYWORDS

formulaic sequence, productive processing, holistic processing, learner-internal formulaic sequence, language proficiency, topic familiarity

1 Introduction

Formulaic sequences (FSs), referring to conventionalized and recurrent word combinations such as idioms, collocations and lexical bundles, have attracted extensive attention from a variety of research fields. In recent years, there has been much debate in psycholinguistic research about how FSs are stored and processed by language users (Siyanova-Chanturia, 2015; Kessler et al., 2020). Some research concludes that FSs are processed holistically without the involvement of grammatical analysis, as they have been found to be processed faster than non-formulaic language (e.g., Underwood et al., 2004; Jiang and Nekrasova, 2007; Millar, 2011; Tremblay et al., 2011; Kim and Kim, 2012; Hallin and Van Lancker Sidtis, 2017). Nevertheless, a growing number of studies have demonstrated that FSs are stored with "live" internal syntactic structures and undergo the same regular syntactic analysis as non-formulaic language, thus suggesting (partial) compositionality of FSs (e.g., Holsinger, 2013; Kyriacou et al., 2020; Mancuso et al., 2020). It is also noted that the above-mentioned studies have mostly examined the receptive processing of FSs. By comparison, the processing of FSs in production tasks remains largely underexplored, particularly in second language (L2) contexts (Siyanova-Chanturia and Martinez, 2015; Siyanova-Chanturia and Lin, 2017). This is regrettable, since research into how L2 learners produce FSs can have important implications for theories concerning the production of L2 FSs as well as the teaching and learning of L2 FSs. The current study therefore aimed to investigate the (conscious) productive processing of FSs by L2 learners in writing.

In addition, there is a need to study FSs from L2 learners' perspective. To date, studies investigating formulaicity in L2 have mostly defined FSs according to native-speaker norms such as authoritative dictionaries and corpus-derived measures, and examined how these idiomatic expressions are used and processed by L2 learners (Myles and Cordier, 2017). Nevertheless, there exists a potential paradox: the targeted FSs might not be known or familiar to L2 learners, thus not serving as holistic, formulaic units for them at all (Schmitt et al., 2004). This issue can be clarified by the important distinction between speaker-external and speaker-internal approaches to formulaicity (Wray, 2008). Speaker-external approaches study conventionalized expressions in the language outside the speaker, identified by external criteria such as formal properties and corpus frequency. Contrastively, speaker-internal or psychological approaches focus on sequences considered formulaic because they are psycholinguistic units for a particular speaker. Underscoring the speaker-internal approaches, this study distinguishes between internal and purely external FSs as the primary criterion of categorizing learners' processing behaviors.

2 Literature review

This section reviews the definitions of FSs and their subtypes, the "holistic or compositional" debate on FS processing, and previous studies on the FS processing types in the learners' production tasks.

2.1 Defining FSs

The most often-cited psycholinguistic definition of FSs was proposed by Wray (2002):

> "a sequence, continuous or discontinuous, of words or other elements, which is, or appears to be, prefabricated: that is, stored and retrieved whole from memory at the time of use, rather than being subject to generation or analysis by the language grammar" (p. 9).

This definition characterizes the holistic property of FSs. However, as Cordier (2013) noted, it seems to suffer a self-contradiction: if the sequence is a discontinuous, flexible formulaic frame with slots for insertion, "it is difficult to conceive that no grammatical processing is taking place at all" (p. 20).

Concerning the identification of psychological FSs, Wray (2008) proposed 11 diagnostic criteria, including previous encounter with the precise formulation, which concerns the speaker's acquisition experience of the FS. Moreover, Myles and Cordier (2017) maintained that a psychological FS should have a holistic quality: "semantic/functional unity or holistic mode of acquisition" (p. 20). The latter means that sequences can receive holistic status, if they are learned as wholes by learners. As can be seen, both Wray (2008) and Myles and Cordier (2017) considered learner's previous acquisition experience of the FS as an important identification criterion.

Importantly, following Wray (2008) and Myles and Cordier (2017) called for a clear awareness of the difference between learner-internal and learner-external FSs. They posited that although there is considerable overlap between what is formulaic for a particular speaker and what is formulaic in the language around this speaker, these two constructs represent different phenomena, and should be investigated as such.

This study follows Wray's (2002) convention to use FS as a coverall for not only psycholinguistic units but also sequences considered formulaic according to external criteria. To capture also external FSs, this study adopts her definition of FSs with modifications: a continuous or discontinuous sequence of words, which appears to be prefabricated, because it is a psycholinguistic unit for a particular learner and/or because it is a conventionalized expression in the language.

Furthermore, separate definitions have been proposed for internal FSs and external FSs. Following Wray (2008) and Myles and Cordier (2017), this study attaches importance to the learner's previous acquisition experience in defining internal FSs. Besides, this study also deems it necessary to establish the holistic status of internal FSs according to the learner's own psychological perception, since formulaicity is viewed as "fundamentally a psychological concept" (Hoey, 2005, p. 7). Therefore, a learner-internal FS is defined as: a continuous or discontinuous sequence of words, acquired previously and perceived as a whole by the learner, rather than being generated word-by-word at the time of use. In this definition, previous acquisition experience of FSs covers not only encountering or learning the FSs from previous linguistic input, but also the fusion of FSs by learners themselves. This is because fusion, which means that previously self-created strings become stored holistically through repeated use, was proposed as an important way of FS acquisition for

L2 learners (Wray, 2002). This definition of learner-internal FSs intends to be exploratory and inclusive. It only claims some degree of holistic representation, which corresponds to some degree of entrenchment (Langacker, 1987; Schmid, 2007; Divjak and Caldwell-Harris, 2015), referring to "the process through which a structure becomes automated into a unit" (Wolter and Gyllstad, 2013, p. 452).

On the other hand, drawing on previous definitions in learner-external approaches of formulaicity (e.g., Qi and Ding, 2011; Jeong and Jiang, 2019; Yu, 2022), this study defines a learner-external FS as: a continuous or discontinuous sequence of words, which has a syntactically and semantically well-formed structure, and can be a conventional way of expressing something.

Additionally, in light of the difference between learner-internal and learner-external FSs (see Figure 1), as emphasized by Myles and Cordier (2017), this study distinguishes between internal and purely external FSs as the primary criterion of categorizing learners' processing behaviors.

As Figure 1 illustrates, although learner-internal and learner-external FSs may overlap considerably (area 2), there would be purely internal FSs (area 1) and purely external FSs (area 3). Purely internal FSs can be seen as idiosyncratic FSs, which are either self-fused strings or low frequency phrases memorized by their users. Such FSs are likely to be neglected by external approaches of formulaicity. Purely external FSs are those conventional FSs that are not perceived as wholes by learners either because of their high compositionality or their low or zero occurrence in the learners' previous linguistic input. These FSs are isolated and termed as assembled FSs in this study, as they are perceived as being assembled word-by-word by the learners. This study does not distinguish between purely internal FSs (area 1) and overlap FSs (area 2) within internal FSs, since they are perceived as wholes indiscriminately from the learners' perspective.

2.2 The "holistic or compositional" debate on FS processing

In psycholinguistic studies, there has been much debate on whether FSs are processed holistically or in a compositional manner. The holistic accounts see FSs as "long words" that are stored and processed holistically, assuming that the components of FSs are not analyzed and there would be no grammatical analysis during their use (e.g., Bobrow and Bell, 1973; Swinney and Cutler, 1979; Gibbs, 1980; Jackendoff, 2002). This assumption of holistic processing has been typically supported by empirical evidence of greater ease in processing FSs than matched non-formulaic phrases, such as shorter reaction time in grammaticality judgment tasks (e.g., Jiang and Nekrasova, 2007), and faster silent reading and articulation (e.g., Tremblay et al., 2011; Kim and Kim, 2012; Hallin and Van Lancker Sidtis, 2017). It has been claimed that FSs enjoy processing advantage because they can bypass the time-consuming syntactic analysis (Swinney and Cutler, 1979). However, this claim has come under criticism. Some researchers pointed out that the processing advantage of FSs did not indicate holistic storage, since it did not concern the relation between the parts and the whole (Arnon and Snider, 2010; Edmonds, 2014; Siyanova-Chanturia, 2015).

Contrastively, the compositional and hybrid accounts emphasize the compositional nature of FSs. Specifically, evidence shows that the literal meanings of component words can be activated during idiom processing (e.g., Cacciari and Tabossi, 1988; Glucksberg, 1993; Sprenger et al., 2006; Cacciari and Corradini, 2015; Beck and Weber, 2016; van Ginkel and Dijkstra, 2019; Kessler et al., 2020); idioms undergo the same regular syntactic analysis as nonidioms (e.g., Cutting and Bock, 1997; Holsinger, 2013); and frequency information of component words still affects the processing of even highly frequent collocations (Arnon and Cohen Priva, 2014; Wolter and Yamashita, 2018; Öksüz et al., 2020).

Importantly, the hybrid accounts view FSs as being holistic and compositional at the same time: while idioms are represented as wholes on some level of processing, they have their internal structures and can be syntactically analyzed on some other level. The holistic nature is reflected in their conventionality and the observation that they are processed faster and more accurately than non-formulaic controls. The compositional nature is revealed by the fact that some of them are decomposable and transparent. As Cieślicka (2010) put it, the apparent inconsistency in FS processing studies can be best resolved by the hybrid accounts of FS representation.

2.3 The hybrid models of idiom representation

One influential model in the hybrid accounts is Model of the lexicon proposed by Cutting and Bock (1997), as shown in Figure 2. From top to bottom, the model consists of three processing levels: conceptual, lexical-conceptual and lexical-syntactic. Idioms are represented as holistic units on the lexical-conceptual level, each having its own lexical-concept node. Meanwhile, idioms are also composed of single words. Horizontally, the model distinguishes between lexicon and syntax. When the lexical-concept node of an idiom is activated, the activation spreads in two directions: one towards the single lemmas that constitute the idiom (lexicon-oriented); the other towards the syntactic information in the form of phrasal frames (syntax-oriented). As the model embraces the dualistic nature of idioms, it has been labeled as a hybrid model of idiom representation.

Later, Sprenger et al. (2006) proposed a modification of the hybrid model, as shown in Figure 3. Specifically, they introduced a *superlemma*, defined as a representation of the idiom on the lemma level, which is a sublevel of the lexical-syntactic level, *lemma* referring to representation of a word's semantic and syntactic information plus a pointer to the word form (see Levelt, 1989; Roelofs, 1992; Jiang, 2000). The adapted model has been termed as the superlemma model, in which idiom production follows the same rules of competition and selection as single words do. For example, the superlemma *hit the road* might compete with *leave*, since there might be competition among co-activated lemmas for the same concept. Importantly, Sprenger et al.

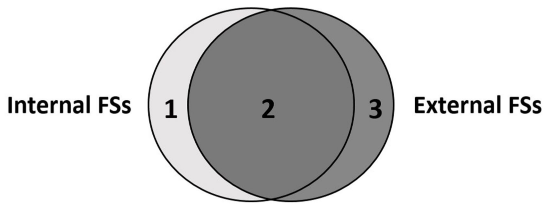


FIGURE 1
The difference between learner-internal and learner-external FSs.

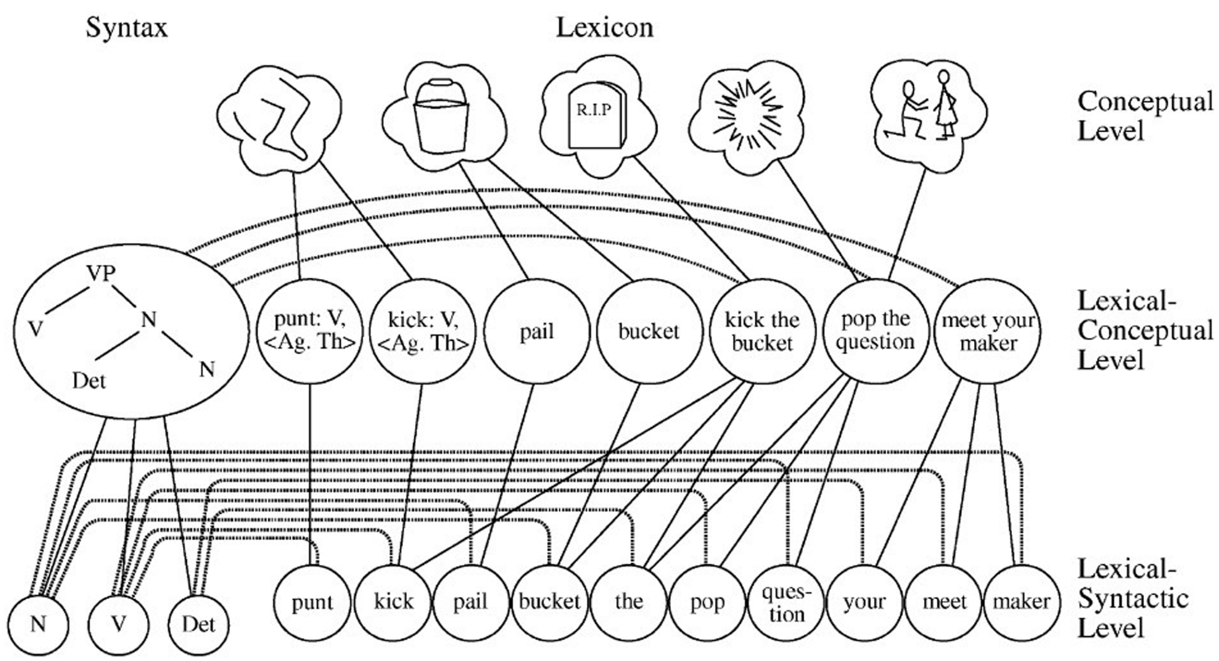


FIGURE 2
Model of the lexicon in Cutting and Bock (1997).

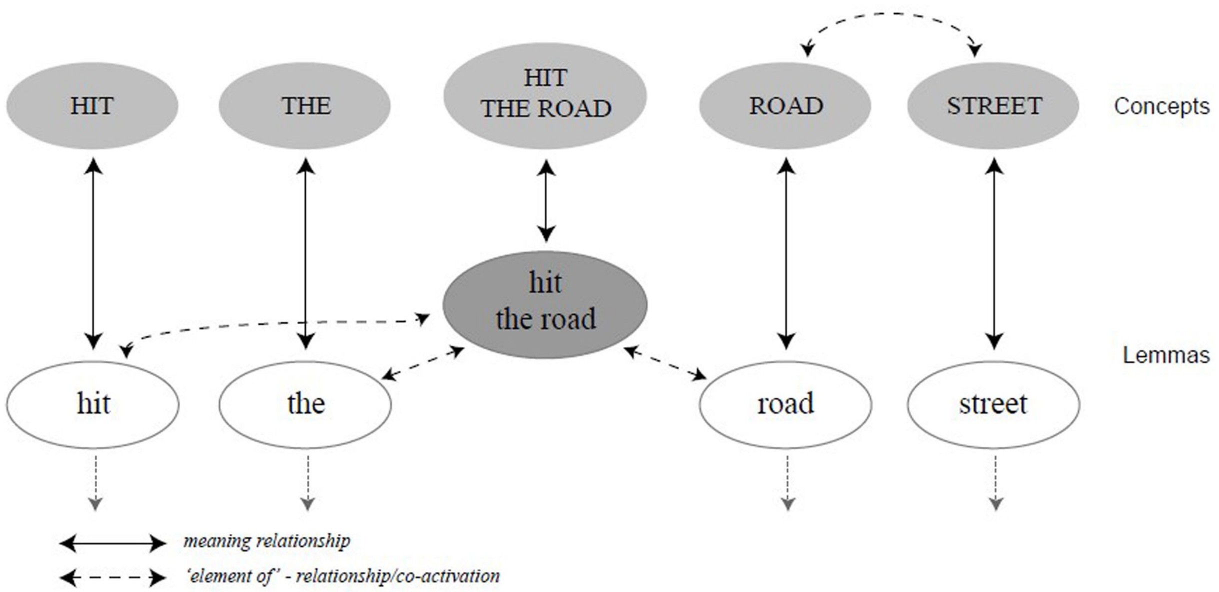


FIGURE 3
The superlemma model in Sprenger et al. (2006).

(2006) contended that the hybrid view can be seen as a general production model of FSs, which may vary in degrees of fixedness, transparency and compositionality.

The hybrid view offers a good solution to the "holistic or compositional" debate on FS processing. Nevertheless, there is still room for improvement. First, the hybrid models are proposed for the production of idioms in the first language (L1). Hence, how they can be adapted for L2 contexts remains to be explored. Second, the basic processing levels in the hybrid models need further specification. The inquiries include such as: What are the chances of co-retrieval of an FS and other lexical items? What if the FS could not be retrieved in its entirety due to inadequate entrenchment in the mental lexicon? How many FSs would be used with syntactic modification? What is the proportion of (conscious) holistic processing on the lexical-syntactic level? Taking these concerns into consideration, the current study examines how FSs are processed by L2 learners in writing. In this way, it could add to the specifications of the lexical-syntactic level in previous hybrid models for L2 contexts.

2.4 FS processing types in the learners' production tasks

So far, only a few studies have investigated the productive processing of FSs. Three of them identified different FS processing types in the learners' production tasks. The earliest one (Spöttl and McCarthy, 2004) explored how multilingual learners processed FSs in a translation task across three or four languages (from L2 into the L1 and then into the third language and/or the fourth language, which can be seen as L2s in a broad sense). Based on think-aloud protocols, three processing types were identified: automatic processing (fluent translation without repetition or evaluation); synthetic evaluative processing (L2 FS repeated and various responses produced and evaluated after a failed attempt at translation); and analytic evaluative processing (FS component words repeated to start the search after a failed attempt at translation). The first two types were described as holistic processing. It was found that synthetic evaluative processing was the most frequent; and automatic processing was employed only occasionally.

In Xu (2010), the participants were English majors from two proficiency levels, writing about the most unforgettable experience in their life. Based on think-aloud protocols, three FS retrieval types were proposed: automatic retrieval (smooth flow of thought), "tip-of-the-tongue" (failed attempt to retrieve the complete form), and piecemeal construction (step-by-step retrieval). Automatic retrieval was found to be the dominant type of retrieval, occupying 81 and 94% for the two groups. Furthermore, higher-proficiency learners had significantly more automatic retrieval and less piecemeal construction than their lower-proficiency counterparts.

Using computer keystroke recordings, Yuan and Xu (2016) investigated the production of FSs by university students in their argumentative L2 writing, and identified three processing types: automatic processing (fluent and fast-rate production, described as holistic processing), semi-automatic processing (fluent but slow-rate production), and controlled processing (dis-fluent production). The authors noted that automatic processing only accounted for 41.98%, while controlled processing was the most frequent. This result differed from Xu's (2010) finding that automatic retrieval was the most frequent, a possible reason being that, compared with the participants in Yuan and Xu (2016), those in Xu (2010) probably wrote on a more familiar topic, thus retrieving more FSs automatically. Topic familiarity has been shown to affect various aspects of learners' production performance, such as fluency (Bui, 2014), lexical complexity (Yang and Kim, 2020; Bui, 2021) and density of FS use (Cordier, 2013). Drawing on Levelt's (1989) model, Bui (2014) suggested that topic familiarity may affect processing at both the Conceptualization and the Formulation stages, resulting in faster access to familiar information and faster retrieval of memorized chunks, which are crucial for the production of FSs.

To summarize, previous studies have proposed tripartite categorizations of FS processing types in learners' production tasks, thus shedding light on the productive processing of L2 FSs. Nevertheless, there are still unresolved issues. First, the processing types involved in L2 FS production await further investigation, as previous studies have yielded inconsistent findings with respect to the proportion of automatic processing or retrieval in L2 FS production, and none of them classified systematically the processing types into holistic or compositional processing. Furthermore, in those studies, the FSs in the learners' language production were identified on the basis of the researchers' judgments, so they might not necessarily match the holistic units from the learners' perspective. Besides, those studies relied on think-aloud method or computer recordings, which may have problems: performing think-aloud might interfere with the normal thinking process (Stratman and Hamp-Lyons, 1994; Sasaki, 2000); the computer-recorded typing might not necessarily mirror the participants' mental activity.

Second, it would be worth exploring the influence of learners' L2 proficiency on their productive processing of FSs. Xu (2010) has suggested the positive relationship between the degree of automaticity in FS production and proficiency development. Besides, the inconsistency in previous findings might stem from differences in the language proficiency of the participants.

Third, another issue concerns the influence of topic familiarity on learners' productive processing of FSs, as is indicated by the difference between the findings in Yuan and Xu (2016) and Xu (2010).

To address the foregoing unresolved issues, this study investigated L2 learners' productive processing of FSs from their own perspective, using the method of video stimulated recall (VSR). While doing so, it also took into consideration the possible effects of L2 proficiency and topic familiarity. The research questions of this study include:

1. How do L2 learners process FSs in writing?

1a. What are the major FS processing types and their frequency/proportion?

1b. Which FS processing types can be seen as holistic processing or compositional processing?

2. What are the effects of L2 proficiency on the learners' productive processing of FSs?

3. What are the effects of topic familiarity on the learners' productive processing of FSs?

Based on the review of previous processing models and empirical studies, the following hypotheses were proposed.

> *Hypothesis 1*: The learners' FS processing types could be categorized on the lexical and syntactic levels (Cutting and Bock, 1997; Sprenger et al., 2006; Xu and Ding, 2010). There would be more holistic processing than compositional processing on each level, as the production of FSs appeared automatic and effortless in most cases according to the learners' verbal reports (Xu, 2010).
>
> *Hypothesis 2*: Higher L2 proficiency would lead to more holistic processing on each level, as L2 proficiency is positively associated with the degree of automaticity in FS production (Xu, 2010).
>
> *Hypothesis 3*: Higher topic familiarity would lead to more holistic processing on each level, as topic familiarity is associated with better performance in learners' production tasks (He and Shi, 2012; Bui, 2014).

3 Methodology

3.1 Design

The current study was primarily qualitative, supplemented by quantitative analysis. It involved collecting and analyzing qualitative data in the first phase of the study and analyzing quantitative data in the second phase. Qualitative analysis was used to delineate learners' processing types while quantitative analysis was conducted to detect if there was any significant difference between the two proficiency groups, or between the two writing tasks. Table 1 provides an overview of the research design. A detailed account follows thereafter.

3.2 Participants

The participants in this study were ten first-year undergraduate students (L1 Chinese, 8 males and 2 females, aged between 18 and 19) at their second semester from a university in East China. At the beginning of their first semester, they all took a comprehensive placement test to be enrolled in a three-level English program for non-English majors (Level 1 presenting the highest proficiency in this population). The English courses for Level-1 students were *College English Reading and Writing* (Level-1) and *College English Listening and Speaking* (Level-1), each having 2 hours of instruction per week. Similarly, Level-3 students attended these two types of courses for Level 3, with the same hours of instruction. Five participants were from Level 1 (hereafter HS1 to HS5 for the five higher-proficiency students), and the other five from Level 3 (hereafter LS1 to LS5 for the five lower-proficiency students). Just prior to the experiment, they all took the College English Test-Band 4, which is a nationwide standardized proficiency test in China for college students. The five higher-proficiency students received an average score of 637.8 (SD = 23.7) out of 710, while the five lower-proficiency students received an average score of 554.8 (SD = 43.1). Therefore, they can be deemed as advanced level and upper intermediate level respectively, as students scoring above 530 were considered as upper-intermediate and advanced (Kessler et al., 2021). Their teachers judged the students' proficiency to approximate level C1 (lower-advanced) and level B2 (upper-intermediate) of the Common European Framework of Reference (CEFR), respectively.

In addition, all of the participants had been raised in China and none had the experience of living abroad. They started learning English from the first or the third grade at primary school (average starting age: 7 (SD = 1.22) for higher-proficiency students and 7.4 (SD = 1.52) for lower-proficiency students; average years of formal instruction: 12.2 (SD = 1.3) for higher-proficiency students and 11.8 (SD = 1.3) for lower-proficiency students). These participants were recruited through random invitation with the help of their teachers and received stationery as gifts for their participation.

3.3 Instruments

3.3.1 Writing tasks

Two argumentative writing tasks differing in topic familiarity were used in this study. The argumentative essay was chosen because it is probably the most common genre practiced at the undergraduate level (e.g., Jiang, 2015; Chen, 2019; Shin, 2019). First, 11 topics were selected from a large number of writing tasks in English tests commonly taken by college students, such as Test of English as a Foreign Language (TOEFL) and The International English Language Testing System (IELTS). Then, 16 freshmen completed a questionnaire to evaluate the familiarity of each topic on a 7-point Likert scale. Half of them were Level-1 students from a parallel class as the HS participants, while the other half were Level-3 students from a parallel class as the LS participants. Thus, they can be seen as representative of the participants in this study. According to the survey results, two topics were chosen (familiarity scores: 5.75 versus 3.38). In both tasks, the participants were required to write about 150–200 words within 35 min. The writing prompts are as follows (written in Chinese in the experiment to avoid any text borrowing):

1. In this fast-paced age, people often confront various kinds of pressure, and college students are no exception. Please write a short essay of 150 to 200 words discussing the reasons for college students' pressure and the solution to it.
2. With the development of economic globalization, the global competition has become more and more fierce. Some people suggest that, to protect the national economy, we should encourage the purchase of domestic products, and limit the purchase of foreign products. How do you view this suggestion? Please write a short essay of 150 to 200 words discussing your viewpoint and giving your reasons.

TABLE 1 Overview of the research design.

Participants	10 freshmen of two L2 proficiency levels
Instruments	Two writing tasks differing in topic familiarity; Training material for students' FS identification; Video stimulated recall (VSR)
Data collection	Session 1: the student writing on the familiar topic→ the training of FS identification→ the student and the researcher identifying FSs separately→ the student and the researcher comparing the two versions to locate potential internal FSs and assembled FSs → the training of VSR → the student performing VSR
	Session 2: the student writing on the unfamiliar topic (The same procedure was repeated except the trainings)
	Session 3: the interview about FS acquisition experience
Data preparation	Transcribing verbal recordings→ ascertaining internal FSs and assembled FSs
Data coding and analysis	Coding of FS processing types (Qualitative analysis) → tallying the descriptive statistics→ non-parametric tests to examine the effects of proficiency and topic familiarity (Quantitative analysis)

3.3.2 Training material for students' FS identification

The current study relied on students' judgments in the identification of learner-internal FSs. After the first writing task, the training of FS identification ensued. A three-part PowerPoint presentation was given to the students in Chinese (see the presentation in Supplementary material). The first two parts present a definition and a categorization of FSs in easy-to-understand language, supplemented by specific examples. The third part presents the identification criteria: at least two words in length; previously learned, encountered or used (including previously self-created expressions which were frequently used later); and being perceived as a whole (having impression of the holistic form, rather than assembling the expression word-by-word on the spot). Students' identification of FSs can serve as a useful way to explore what is formulaic in their mind, since it has been found that laypeople's intuitive judgments of formulaicity are valid (Wulff, 2008; Lin, 2018), and L2 learners' intuitions are a reliable predictor of their idiom knowledge (Hubers et al., 2020).

3.3.3 Video stimulated recall (VSR)

VSR interview is a technique for investigating the participants' cognitive processes by promoting them to recall their thinking while playing video-recordings of their own behaviors (Dempsey, 2010). This technique has been proved highly effective in process-oriented writing studies (e.g., Abdi Tabari, 2022). It is suggested that VSR be conducted as soon as practicable to prevent recall failure (Dempsey, 2010; Gass and Mackey, 2017). Upon the completion of the writing and the FS identification (as will be illustrated below), the participants were provided with instructions and demonstrations on how to perform VSR. Then they performed VSR while watching the video-recordings of their writing process captured by the screen recording software Camtasia Studio.[1] During the recall, the video could be paused or played back to allow for detailed explications. The participants were prompted by questions such as "What were you thinking at that moment?" "Why did you make this change?" "Why did you pause here?" and so on. Besides, after the recall of each paragraph, the student also recalled the production process of each FS according to the researcher's prompt questions (see the prompt questions in Supplementary material).

3.4 Data collection

Data collection was carried out in three sessions for each participant on a one-to-one basis. In the first session, the student wrote about the familiar topic in a Word file (without auxiliary functions) on a computer. Meanwhile, the software Camtasia Studio was used to record the writing process on the screen. The next was the training of FS identification. After that, the student and the researcher identified the FSs, individually on separate computers. The student was required to mark in red the expressions they considered as FSs, while the researcher identified potential external FSs according to the criterion of structural completeness and semantic unity. Then, collectively, comparison was made between the two annotated versions. Sequences identified by the researcher but not the student were rechecked immediately, and then marked in different colors by the student: those acknowledged as FSs by the student were marked red since they had been missed simply because of overlooking, while those that the student claimed as being assembled from scratch were marked blue. Consequently, the "red" sequences were potential internal FSs to the student, while the "blue" sequences (researcher-only sequences) were potential assembled (purely external) FSs to him or her. Then, the student performed VSR for the FS production processes.

Three days later, in the second session, the writing topic changed to an unfamiliar one. The same procedure was repeated, except the training of FS identification and VSR. One day or several days thereafter, in the third session, the students were interviewed about their acquisition experience of the FSs they had identified. The interview began with some questions concerning the student's FS acquisition experience in general, and then the acquisition experience of each FS was inquired with a series of questions (see the interview guide in Supplementary material).

3.5 Data preparation

All the verbal recordings were transcribed by the first author with the help of the software Iflyrec.[2] As a next step, the potential internal FSs and assembled FSs needed to be ascertained for their formulaic status.

3.5.1 Ascertaining internal FSs

The interview data about FS acquisition experience were used to ascertain the formulaic status of internal FSs, that is, to check if the FSs identified by the students had been acquired previously as wholes by them. Based on the interview data, this study classified the students' FS acquisition experience into four categories: 1. deliberate memorization; 2. incidental learning without a particular intention to learn; 3. brief noticing or semi-intentional learning; and 4. fusion (acquiring the FS through self-construction and later frequent use).[3] The formulaic status of an internal FS was ascertained if the student's acquisition experience of it belonged to one of the four categories. Note that the internal FSs might be erroneous (e.g., **handle with*), as such erroneous forms probably had been entrenched in the students' mental lexicon by repeated use.

3.5.2 Ascertaining assembled FSs

In identifying external FSs, we followed the FS identification method in Qi and Ding (2011) with some modifications. The criteria are as follows: 1. being composed of two or more than two words and having structural completeness and semantic unity; 2. being contained

1 https://www.techsmith.com/download/camtasia/

2 https://www.iflyrec.com/

3 We found that fusion accounted for 4%, e.g., *cast a significant influence on* (HS3: Perhaps this was created by myself. I learned that "cast" means 投掷, like in "cast shadow." I thought it might have a more abstract meaning, so I used "cast a significant influence on." I have used it frequently in writing since high school).

in the *Longman Dictionary of Contemporary English*[4] or *Oxford Collocation Dictionary of English*[5]; or being listed as a collocation or cluster in the Corpus of Contemporary American English; 3. being confirmed by native speaker intuition.

During the FS identification immediately after each writing, the first criterion was applied by the researcher to locate external FSs. To ascertain the formulaic status of potential assembled FSs, the second and the third criteria needed to be applied to determine whether they qualified as authentic FSs in the language. Note that if the sequence resembled the standard form considerably, it was treated as an inappropriately assembled FS (e.g., **life quality/quality of life*). The first author and a research assistant checked separately the potential assembled FSs in the dictionaries and the corpus. The differences in the results were settled through discussion and a unanimous list of FSs was reached. Finally, the FS list was presented to a native speaker for further confirmation. It turned out that all the FSs on the list were confirmed by native speaker intuition, as they had been carefully checked against authoritative sources including the dictionaries and the corpus.

3.6 Data coding and analysis

Based on the students' VSRs and computer recordings, iterative analysis was carried out to categorize the students' processing behaviors during their FS production. In line with previous models of language representation and production (Cutting and Bock, 1997; Sprenger et al., 2006; Xu and Ding, 2010), two processing levels were distinguished hierarchically: lexical and syntactic. These two levels can be seen as sub-levels of the lexical-syntactic level in Cutting and Bock (1997).

At the lexical level, the FS basic form is retrieved or assembled by the student. It is similar to the lexical retrieval stage in Xu and Ding (2010), where writers make efforts to retrieve the lexical items needed to convey the intended meaning. In the current study, categorization on this level primarily draws upon the distinction between internal FSs and assembled FSs, with the former being retrieved on the basis of the student's impression of the holistic phrasal forms, while the latter being constructed word-by-word without a holistic base and coinciding with a conventionalized expression in the language. Then, within internal FSs, further categorization is made according to the literature (Xu, 2010) and the data of the present research. It is noteworthy that "the basic form" is conceptualized differently for internal and assembled FSs. For the former, it refers to the holistic form of the FS in the student's mind, that is, the form that the student retrieves as a whole in the first place. These basic forms are delineated according to the students' description of the phrasal form that appeared in their mind first for a certain meaning. For assembled FSs, the basic form refers to the student's combination of individual words that coincides with an authoritative expression in the dictionaries and the corpus, such as *feel nervous*. In a sense, the FS basic form resembles superlemma in Sprenger et al. (2006). Nevertheless, it is the basic form of a holistically acquired phrase or a word combination from the student's perspective, rather than in linguistic terms. Actually, the FS basic forms in this study might not be linguistically lemmatized forms as would appear at the beginning of a dictionary entry. For example, *ranging from…to* and *are easily to* were retrieved as FS basic forms.

At the syntactic level, the FS basic form is embedded in the text either intact (intact integration) or with modification (syntactic modification) by the student. It is similar to the formal integration stage (Xu and Ding, 2010), where writers attend to the formal features of lexical items and embed them in specific contexts. In the current study, intact integration is identified if the FS basic form remains the same in the written product; syntactic modification is detected according to the student's description of the modification they made to the FS basic form in their conscious mind. For internal FSs, the unit of modification is the entire FS, and the modification can be made by lexical and morphological means (e.g., *be different from→is greatly different from*). For assembled FSs, the unit of modification is the individual component words, so the modification can only happen by morphological means (e.g., *feel nervous→feels nervous*).

After the coding scheme had been developed, the first author and a research assistant coded separately four randomly selected VSRs, and the inter-coder reliability was 0.92. The differences were resolved through discussion. Then, the first researcher coded the remaining data.

After the coding was done, the frequency and percentage of each FS processing type in each composition were tallied. Considering the small sample size, the current study conducted non-parametric tests using SPSS 25 for the quantitative analysis. The two independent variables include one between-participant variable (two proficiency levels) and one within-participant variable (familiar and unfamiliar topics). The dependent variable—learners' FS processing—was measures in terms of frequency and percentage of FS processing types. Specifically, based on the descriptive data, Mann–Whitney U tests (Two-independent samples tests) were run to detect variances between the two proficiency groups, and Wilcoxon Signed Ranks tests (Two-related samples tests) to detect variances between the two writing tasks.

4 Results

4.1 Overall description of L2 learners' processing of FSs in writing

4.1.1 Major FS processing types and their frequency/proportion

The learners' FS processing types are categorized on two processing levels: lexical and syntactic. On the lexical level, drawing on Xu's (2010) taxonomy of FS retrieval types, the current study identified five FS processing types, i.e., single retrieval, parallel retrieval, part-to-whole retrieval, "din in the head" and online assembly. The first four types describe the processing of internal FSs, while online assembly denotes the processing of assembled FSs.

Single retrieval means that the FS is retrieved fluently in its entirety as the single choice for a certain meaning. For example:

> Upon seeing the writing topic, I felt that it is about a common phenomenon. So I came up with a chunk "it is universally acknowledged that." (HS5-Familiar)

4 https://www.ldoceonline.com/
5 http://www.freecollocation.com/

Parallel retrieval means that the FS is retrieved simultaneously with other expression (s) for the same meaning. For example:

> For 各种各样的压力 [all kinds of pressure], several expressions flashed up in my mind, like "various" and "a variety of." I used "all kinds of" because it was more familiar to me. (HS1-Familiar)

Part-to-whole retrieval means that the FS is retrieved not in its entirety, but rather in a part-to-whole manner. Specifically, the writer retrieved a part of the FS first, and then retrieved the remaining part either fluently or laboriously. This indicated that the component words of FSs do not always have equal weighting, with some being more salient and more easily retrievable than others. In the following example, *take charge of* was retrieved by laborious part-to-whole retrieval, as the student had struggled to recall the final word of it:

> college students need to take charge of themselves.
>
> I was hesitating between "for" and "of" for quite a while. (HS4-Familiar)

"Din in the head" means that the target FS cannot be successfully retrieved at the moment despite the student's retrieving effort. To be specific, the students had an ideal FS in their mind for the current use, but they were unable to retrieve its form successfully. The term "din in the head" was originally defined as "the sense of having the language available for use" (Krashen, 1985, cited from Cohen, 1998, p. 244). This term, rather than "tip of the tongue" (Xu, 2010), is used in the current study, as it implies only a weak memory trace of the expression. Indeed, students may come across the disappointing situation that the memory trace of the desired FS was too weak that they failed to retrieve its form. "Din in the head" differs from part-to-whole retrieval in that it denotes failed retrieval, though they both entail construction efforts. For example:

> I wanted to express 分轻重缓急 [get your priorities right] and thought of a newly learned chunk for this meaning, but I couldn't recall it, so I gave up. (HS2-Familiar)

Online assembly means that the FS is assembled word-by-word on the spot. These FSs are the researcher-only FSs, i.e., only identified by the researcher and deemed as being assembled or improvised by the students. Online assembly can also be fluent or laborious, depending on whether the student had difficulties during the assembling process. Notably, online assembly differs from part-to-whole retrieval and "din in the head" in that it denotes word-by-word construction from scratch, while the other two denote construction on the basis of some vague or "worn-out" memory traces. Consider an example of fluent online assembly:

> (*reduce pressure*) To express 减轻压力, I thought of 压力 [pressure] first, and then 减轻 [reduce]. Then I judged whether they could collocate. As I thought they could, I put them together. (LS4-Familiar)

Among the total 45.95 FSs retrieved or assembled during a writing task averagely, single retrieval was the most frequent (mean frequency/percentage = 27.05/58.16%), followed by online assembly (10.20/23.32%), parallel retrieval (5.55/11.44%), part-to-whole retrieval (2.70/6.09%) and "din in the head" (0.45/0.99%). Furthermore, the accumulated frequency/percentage of internal FSs (all the categories except online assembly) is 35.75/76.68%.

On the syntactic level, a distinction was made between intact integration and syntactic modification, depending on whether the FS basic forms were used intact or with modification. Intact integration means that the FS basic form is embedded intact in the text without any modification. For instance:

> "A good choice" came out directly, and I made no change to it. Ah, why didn't I think over "good"? "Good" can certainly be changed for a better word. What a pity. (LS4-Familiar)

Interestingly, it was found that the FS basic forms are not necessarily linguistically lemmatized forms. Rather, they might contain inflected words. For example:

> I always use "we are supposed to." Lots of writings are about suggestions. Although the collocation is "be supposed to," I commonly use "are supposed to" directly, and seldom use "be supposed to." (LS2-Familiar)

As illustrated above, the FS *we are supposed to*, fully specified in grammatical features, was embedded intact in the text. This suggests that such grammatical markers may have been frequently used with the particular FS to the extent that they have become an integral part of it. Either the inflected forms have also been stored in the mental lexicon, or the syntactic-morphological operations have become so automatized that they do not need any conscious effort. This is consistent with the hypothesis that FSs may be stored at different levels of abstraction (Ambridge and Lieven, 2011; Cordier, 2013; Wulff, 2018). For example, Cordier (2013) found that while learners seem to store abstract formulaic frames, they may also have automatized some fixed, specific sequences.

On the other hand, syntactic modification means that the FS basic form is modified in one way or another according to the specific context. It was found that learners' syntactic modification could happen in the morphological aspect such as person, tense, participle, and determiner, or in the lexical aspect including addition, substitution and omission of words within the FS, or in both. For example:

Morphological modification (participle):

> a "native" product may *has its raw materials originating from other* countries.
>
> I thought of "originate from." I knew participle should be used, yet hesitating between present participle and past participle. (HS3-Unfamiliar)

Lexical modification (substitution):

> *I'm appreciate to share some opinions about it with you.
>
> This is a frequently used sentence pattern. I thought "glad" was quite clichéd, so I substituted it with "appreciate." (LS1-Unfamiliar)

Among the 41.45 FSs embedded per text on average, intact integration (mean frequency/percentage = 27.25/66.53%) was much more frequent than syntactic modification (14.20/33.47%). Note that the frequency of embedded FSs was lower than that of retrieved/assembled FSs (45.95). This is because some FSs, albeit being retrieved or assembled on the lexical level, were obsoleted without being embedded, thus failing to reach the syntactic level. Furthermore, internal FSs with intact integration were the most frequent (mean frequency/percentage = 19.55/46.79%), followed by internal FSs with syntactic modification (11.75/27.78%), assembled FSs with intact integration (7.70/19.7%), and finally assembled FSs with syntactic modification (2.45/5.69%).

4.1.2 Dividing the FS processing types into holistic or compositional processing

In this study, holistic processing is conceptualized as the processing that does not involve writers' conscious, overt syntactic analysis of the FSs into component words. In the two-layered categorization, on the lexical level, single retrieval and parallel retrieval can be seen as holistic processing in a sense. By contrast, the other types should be regarded as compositional processing, since they all entail some constructional efforts. On the syntactic level, intact integration of internal FSs can be seen as holistic processing in a sense (though there might be some minimal syntactic analysis), while syntactic modification of internal FSs should be regarded as compositional processing, since it is the overt manifestation of syntactic analysis. Besides, since assembled FSs were not perceived as holistic units by the learners, all the integration of assembled FSs belongs to compositional processing. Note that the difference between holistic processing and compositional processing is a matter of degree, as holistic processing is a gradable concept (Boers et al., 2006) and idiomaticity is a scalar property (Wulff, 2008).

Consequently, the results of the present study can give an indication of the relative proportion of holistic versus compositional processing on each level. On the lexical level, holistic processing may account for 69.60% (the accumulated percentage of single and parallel retrieval), which is essentially determined by the percentage of internal FSs (76.68%), since internal FSs retrieved through compositional processing (part-to-whole retrieval and "din in the head") are quite scarce. On the syntactic level, holistic processing may account for 46.79% (the percentage of intact integration of internal FSs). Inversely, compositional processing may account for 30.40% on the lexical level and 53.21% on the syntactic level.

4.2 The productive processing of FSs by the two proficiency groups

To answer the second research question, comparisons were made between the two proficiency groups with respect to the productive processing of FSs. Table 2 presents the descriptive results and the statistical test results concerning the between-group comparison of the five FS processing types on the lexical level. Mann–Whitney U tests showed that higher-proficiency students had significantly more single retrieval ($Z = -2.656$, $p < 0.01$) and parallel retrieval ($Z = -2.621$, $p < 0.01$) in frequency, while lower-proficiency students had significantly more online assembly in percentage ($Z = -1.978$, $p < 0.05$). Furthermore, the accumulated frequency/percentage of internal FSs was 42.8/81.76% for higher-proficiency students and 28.7/71.59% for lower-proficiency students, while the frequency/percentage of assembled FSs was 9.4/18.24% for higher-proficiency students and 11.0/28.41% for lower-proficiency students. This indicates that compared with lower-proficiency students, the higher-proficiency group were more likely to retrieve prefabricated expressions from their mental lexicon, rather than assemble FSs from scratch.

Table 3 reports the descriptive results and the statistical test results concerning the between-group comparison of the two FS processing types on the syntactic level. Mann–Whitney U tests revealed that the two groups had significant differences on this level: higher-proficiency students employed significantly less intact integration in percentage ($Z = -2.960$, $p < 0.01$), and made significantly more syntactic modifications in frequency ($Z = -3.600$, $p < 0.001$) and percentage ($Z = -2.960$, $p < 0.01$). Table 4 further demonstrates that these differences mainly lie in the processing of internal FSs: higher-proficiency students made significantly more modifications to internal FSs (Frequency, $Z = -3.413$, $p < 0.01$; Percentage, $Z = -2.613$, $p < 0.01$).

Furthermore, on the lexical level, the proportion of holistic processing was 74.93% for higher-proficiency students and 64.27% for lower-proficiency students. This difference resulted from higher-proficiency students' retrieval of more internal FSs than lower-proficiency students. On the syntactic level, the proportion of holistic processing was 45.07% for higher-proficiency students and 48.51% for lower-proficiency students. Despite retrieving more internal FSs, higher-proficiency students nevertheless made more modifications to them, thus reducing holistic processing on the syntactic level. Taken together, the results show that higher L2 proficiency may lead to higher proportion of holistic processing on the lexical level but not the syntactic level.

4.3 The productive processing of FSs in the two writing tasks

To answer the third research question, comparisons were made between the two writing tasks with respect to the productive processing of FSs. Table 5 presents the descriptive results and the statistical test results concerning the between-task comparison of the five FS processing types on the lexical level. It can be seen that the two tasks resembled each other considerably on this level. Wilcoxon Signed Ranks tests showed that the only significant difference was the higher frequency of single retrieval in the familiar-topic writing ($Z = -2.524$, $p < 0.05$). This contributes to the higher frequency of internal FSs in the familiar-topic writing than in the unfamiliar-topic writing (39.5 versus 32). Nevertheless, the proportions of internal FSs in the two tasks were almost the same (77.25% versus 76.12%), since there were also more assembled FSs in the familiar-topic writing (10.8 versus 9.6).

Table 6 reports the descriptive results and the statistical test results concerning the between-task comparison of the two FS processing types on the syntactic level. It shows that the two tasks were strikingly similar in terms of the two broad categories on this level: the category percentages were almost the same across the two tasks. Wilcoxon Signed Ranks tests showed no significant difference in the two broad categories. Despite the overall similarity on the syntactic level, Table 7 reveals a more complex picture: for internal FSs, the learners made

TABLE 2 Between-group comparison on the lexical level.

Categories	Frequency				Percentage (%)			
	High	Low	Z	Sig.	High	Low	Z	Sig.
Single retrieval	31.6 (7.4)	22.5 (6.5)	−2.656	0.008**	60.02 (5.98)	56.30 (11.33)	−0.643	0.520
Parallel retrieval	7.9 (4.8)	3.2 (2.3)	−2.621	0.009**	14.91 (7.93)	7.97 (5.90)	−2.015	0.044*
Part-to-whole retrieval	2.7 (2.1)	2.7 (1.8)	−0.193	0.847	5.64 (5.40)	6.53 (3.95)	−0.874	0.382
Din in the head	0.6 (0.8)	0.3 (0.5)	−0.717	0.473	1.19 (1.64)	0.79 (1.29)	−0.534	0.593
Online assembly	9.4 (2.7)	11.0 (3.5)	−1.031	0.303	18.24 (4.72)	28.41 (10.74)	−1.978	0.048*

High, Higher-proficiency group; Low, Lower-proficiency group; Standard deviation is given in parentheses. *, significant at $p < 0.05$; **, significant at $p < 0.01$.

TABLE 3 Between-group comparison on the syntactic level.

Categories	Frequency				Percentage (%)			
	High	Low	Z	Sig.	High	Low	Z	Sig.
Intact integration	27.6 (5.7)	26.9 (6.0)	−0.152	0.879	59.85 (8.66)	73.21 (6.34)	−2.960	0.003**
syntactic modification	18.7 (5.7)	9.7 (2.3)	−3.600	0.000***	40.15 (8.66)	26.79 (6.34)	−2.960	0.003**

High, Higher-proficiency group; Low, Lower-proficiency group; Standard deviation is given in parentheses. **, significant at $p < 0.01$; ***, significant at $p < 0.001$.

significantly less modifications to them in the familiar-topic writing (Percentage, $Z = -2.193$, $p < 0.05$), while the reverse was detected for assembled FSs: the learners made significantly less modifications to them in the unfamiliar-topic writing (Frequency, $Z = -2.869$, $p < 0.01$; Percentage, $Z = -2.812$, $p < 0.01$).

Furthermore, on the lexical level, the proportion of holistic processing was 71.12% in the familiar-topic writing and 68.08% in the unfamiliar-topic writing. Such high similarity resulted from the fact that the percentage of internal FSs was about the same in the two tasks. On the syntactic level, the proportion of holistic processing was 50.54% in the familiar-topic writing and 43.04% in the unfamiliar-topic writing. This difference can be attributed to the fact that learners had more intact integration of internal FSs in the familiar-topic writing. Taken together, the results showed that topic familiarity may lead to higher proportion of holistic processing on the syntactic level, but not the lexical level.

5 Discussion

5.1 L2 learners' productive processing of FSs

In line with the first hypothesis, the learner's FS processing types can be categorized on the lexical and syntactic levels. Furthermore, the findings partially support the first hypothesis in that there was more holistic processing than compositional processing on the lexical level, but not on the syntactic level. Specifically, the high proportion of holistic processing on the lexical level was mainly driven by the frequent retrieval of internal FSs, while the reduced proportion of holistic processing on the syntactic level was mainly caused by the substantial amount of modification made to the internal FSs.

5.1.1 FS processing on the lexical level: frequent retrieval of internal FSs

On the lexical level, the retrieval of internal FSs had a frequency of 35.75, accounting for 76.68% of FS processing. This result indicates that the learners frequently retrieved internal FSs from their mental lexicon, given the short length of the writing (about 200 words). This contradicts Wray's (2002) claim that classroom L2 learners tend to store words separately, but supports other previous findings that learners retain information about the co-occurrence of words. Wray (2002) claimed that the creation of L2 lexicon is fundamentally different from that of L1 lexicon. When encountering *major catastrophe*, native speakers would notice and store it as a sequence. In contrast, L2 learners would analyze it into individual words. Consequently, they are likely to acquire a lexicon consisting of single words. However, counterevidence has been found against this claim. For example, repetition can promote the incidental learning of L2 collocations (e.g., Durrant and Schmitt, 2010; Webb et al., 2013); L2 learners are sensitive to the phrasal frequency of FSs (e.g., Ellis et al., 2008; Wolter and Gyllstad, 2013; Wolter and Yamashita, 2018; Northbrook and Conklin, 2019; Öksüz et al., 2020). Along these lines, the current findings suggest that L2 learners have considerable storages of FSs in their mental lexicon.

It can be argued that even though learners may tend to break down word sequences into individual words, they may also pay attention to how the words glue together and memorize the sequences as wholes. In other words, they may attend to individual words and the whole FS simultaneously. Furthermore, learners' formation analysis of FSs could actually facilitate retention, as seen from the interview excerpt:

> (be faced with) This form felt a little strange. It could express a sense of enforcement, as if the fate compels you to face it, like "face somebody something." You are compelled to face it, rather than voluntarily. (HS5)

In this example, the student analyzed the semantic structure of the FS to understand the form-meaning mapping (the reason why the form expresses the meaning), which could potentially help memorization. Actually, the mnemonic benefits of learners' analysis of FSs have been increasingly recognized and verified (e.g., Boers and Lindstromberg, 2009; Hatami, 2015).

TABLE 4 Between-group comparison within internal FSs and assembled FSs on the syntactic level.

Categories		Frequency				Percentage (%)			
		High	Low	Z	Sig.	High	Low	Z	Sig.
Internal FSs	Intact integration	21.1 (6.0)	18.0 (6.1)	−1.138	0.255	45.07 (7.10)	48.51 (9.35)	−1.099	0.272
	Syntactic modification	15.7 (4.0)	7.8 (3.3)	−3.413	0.001**	34.23 (8.45)	21.34 (8.90)	−2.613	0.009**
Assembled FSs	Intact integration	6.5 (3.0)	8.9 (2.6)	−1.638	0.101	14.79 (7.58)	24.70 (7.88)	−2.196	0.028*
	Syntactic modification	3.0 (2.7)	1.9 (1.2)	−0.695	0.487	5.92 (4.30)	5.45 (3.94)	−0.038	0.970

High, Higher-proficiency group; Low, Lower-proficiency group; Standard deviation is given in parentheses. *, significant at $p < 0.05$; **, significant at $p < 0.01$.

TABLE 5 Between-task comparison on the lexical level.

Categories	Frequency				Percentage (%)			
	Familiar	Unfamiliar	Z	Sig.	Familiar	Unfamiliar	Z	Sig.
Single retrieval	30.7 (9.0)	23.4 (5.8)	−2.524	0.012*	60.40 (8.61)	55.92 (9.31)	−1.581	0.114
Parallel retrieval	5.7 (5.1)	5.4 (3.9)	−0.141	0.888	10.72 (8.14)	12.16 (7.56)	−0.833	0.405
Part-to-whole retrieval	2.5 (1.9)	2.9 (2.0)	−0.539	0.590	4.86 (3.55)	7.32 (5.40)	−1.053	0.292
Din in the head	0.6 (0.8)	0.3 (0.5)	−1.000	0.317	1.27 (1.70)	0.72 (1.17)	−0.850	0.395
Online assembly	10.8 (3.0)	9.6 (3.3)	−1.196	0.232	22.76 (9.42)	23.89 (10.28)	−0.416	0.677

Familiar, Familiar-topic writing; Unfamiliar, Unfamiliar-topic writing; Standard deviation is given in parentheses. *, significant at $p < 0.05$.

5.1.2 FS processing on the syntactic level: substantial amount of modification

On the syntactic level, the result that syntactic modification accounted for 33.47% indicates that a substantial proportion of FSs were indeed modified at the time of use. This result converges with previous findings that point to the syntactic flexibility of FSs (e.g., Gibbs and Nayak, 1989; Barkema, 1994; Moon, 1998; Grant, 2003; Wulff, 2008; Kyriacou et al., 2020). For example, Wulff (2008) found that the majority of the targeted idiomatic V NP-constructions did not deviate strongly from the baseline in terms of syntactic flexibility. Kyriacou et al. (2020) showed that nontransparent idioms can be passivized while retaining their figurative meaning. These previous findings point to the fact that FSs are much more flexible than commonly assumed, and the current findings testify to this from a process-based, learner-internal perspective.

Consequently, the implication is that the corpus-driven approaches which did not take syntactic flexibility into account might overlook many genuine instances of FSs and thus seriously underestimate the frequency of flexible FSs (e.g., Biber et al., 1999; Staples et al., 2013; Huang, 2015). For example, Biber et al. (1999) noted that verb phrase bundles were rarely found in academic discourse. Despite register influence, such rarity is possibly because verbs in English are most likely to occur in various, inflected forms which might be missed by the concordancers. By comparison, corpus-driven approaches that used lemmatized frequency counts and allowed distances within collocations can better accommodate the syntactic flexibility of FSs (e.g., Vincent, 2013; Durrant, 2014; Yoon, 2016).

Additionally, what is particularly noteworthy about learners' syntactic modification of FSs is that such modification might lead to errors in FS use. It was found that students might retrieve a correct FS from the mental lexicon, but made inappropriate modification to it. For example, a student modified *suffer from* inappropriately:

> They *are usually suffered from the academic stress.
>
> I changed "suffer from" into passive form. They are tortured by something. These victims should be passive, being scared. (LS3-Familiar)

In this example, an error occurred as the student did not master the usage of the FS adequately. This shows that knowing the basic form of an FS is just the first step, while being able to integrate it appropriately into writing is equally important.

5.1.3 A model of L2 FS production

Based on previous hybrid models of idiom production (e.g., Cutting and Bock, 1997; Sprenger et al., 2006) and the current categorization of FS processing types, a model of L2 FS production was proposed (see Figure 4). This model consists of three processing levels: conceptual, lexical and syntactic. After generating a concept, the learner either retrieves an internal FS (s) from the mental lexicon (which can take the form of single retrieval, parallel retrieval, part-to-whole retrieval and "din in the head"), or assembles an expression from scratch (online assembly). Subsequently, the learner embeds the retrieved or assembled FS in the text through either intact integration or syntactic modification. Concerning the characteristics of the processing types, the ellipse in the figure indicates holistic processing, while the rectangle indicates compositional processing.

By specifying L2 learners' FS processing types on different levels, this tentative model of L2 FS production extends previous hybrid

TABLE 6 Between-task comparison on the syntactic level.

Categories	Frequency				Percentage (%)			
	Familiar	Unfamiliar	*Z*	Sig.	Familiar	Unfamiliar	*Z*	Sig.
Intact integration	29.8 (6.4)	24.7 (3.6)	−1.790	0.074	67.33 (11.10)	65.73 (9.47)	−0.614	0.539
syntactic modification	15.3 (7.8)	13.1 (4.3)	−1.131	0.258	32.67 (11.10)	34.27 (9.47)	−0.614	0.539

Familiar, Familiar-topic writing; Unfamiliar, Unfamiliar-topic writing; Standard deviation is given in parentheses.

TABLE 7 Between-task comparison within internal FSs and assembled FSs on the syntactic level.

Categories		Frequency				Percentage (%)			
		Familiar	Unfamiliar	Z	Sig.	Familiar	Unfamiliar	Z	Sig.
Internal FSs	Intact integration	22.9 (6.6)	16.2 (3.3)	−2.710	0.007**	50.54 (6.46)	43.04 (8.45)	−2.298	0.022*
	Syntactic modification	11.6 (6.2)	11.9 (4.7)	−0.358	0.720	24.63 (10.23)	30.94 (10.77)	−2.193	0.028*
Assembled FSs	Intact integration	6.9 (3.1)	8.5 (2.8)	−1.611	0.107	16.80 (9.43)	22.69 (8.11)	−1.990	0.047*
	Syntactic modification	3.7 (2.3)	1.2 (0.9)	−2.869	0.004**	8.04 (3.78)	3.33 (2.75)	−2.812	0.005**

Familiar, Familiar-topic writing; Unfamiliar, Unfamiliar-topic writing; Standard deviation is given in parentheses. *, significant at $p < 0.05$; **, significant at $p < 0.01$.

models of idiom production, thus contributing to the "holistic or compositional" debate on FS processing. The previous hybrid models were proposed for L1 idiom production, which is supposed to be highly automatized. However, the production of L2 FSs might be a more complex and effortful process, as Jiang (2000) contended that the development of most L2 words would fossilize before reaching the final, native-like stage. In the current model, on the lexical level, part-to-whole retrieval, "din in the head" and online assembly all entail some constructional efforts. For L1 speakers, the FS basic forms are supposed to be retrieved holistically due to deeply entrenched mental representations. However, for L2 learners, their mental representations of even internal FSs may be relatively weak, which is manifested by part-to-whole retrieval and "din in the head." Besides, they do not have unitary representations of the assembled FSs (at least not perceivable), so they have to construct these FSs from single words. On the syntactic level, syntactic modification is proposed as an FS processing type. According to Jiang (2000), whereas the production of morphologically appropriate words tends to be an automatic process for L1 speakers, it is likely to be "a conscious process involving two strictly serial steps: the selection of a root form such as *leave* and then the selection of the morphologically appropriate form" (p. 58). In the current model, these two serial steps occur on the lexical level and the syntactic level, respectively. Therefore, while previous hybrid models only indicate that idioms can be compositional, without denoting any consciousness of modification on the speakers' part, this study establishes syntactic modification as a conscious manipulation by L2 learners.

By elaborating the partial compositionality of FSs on different processing levels, this study can help to interpret the inconsistency between previous findings. Using computer keystroke recordings, Yuan and Xu (2016) found that controlled processing of FSs was most frequent, which denotes dis-fluent production. According to the current model, such dis-fluency might result from the compositional processing on the lexical and syntactic levels. Actually, the percentage of "single/parallel retrieval → intact integration," entailing holistic processing on both levels, was 43.67%, which is strikingly similar to the result of Yuan and Xu (2016) that automatic processing only accounted for 41.98%. By contrast, using think-aloud, Xu (2010) found that automatic retrieval dominates. This might result from the high percentage of fluent processing types on the lexical level (the accumulated percentage of single/parallel retrieval, fluent part-to-whole retrieval and fluent online assembly was 90.37% in the current study), since think-aloud method is capable of revealing the learners' mental processes on this level.

5.2 Effects of L2 proficiency on learners' productive processing of FSs

The findings partially support the second hypothesis in that higher L2 proficiency led to more holistic processing on the lexical but not the syntactic level. Specifically, compared with lower-proficiency students, higher-proficiency students had retrieved more internal FSs and made more syntactic modifications to them.

Higher-proficiency students retrieved more internal FSs probably because they had higher awareness of FS memorization and thus larger FS storage. In the interview, when asked about whether they would intentionally collect and memorize FSs in normal situations, all the higher-proficiency students responded affirmatively, while three out the five lower-proficiency students responded negatively. With higher awareness of FS memorization, higher-proficiency students would not only fulfill conscientiously the FS memorization tasks assigned by their teachers, but also pay attention to the worthy FSs on their own initiative, as seen in the following interview excerpt:

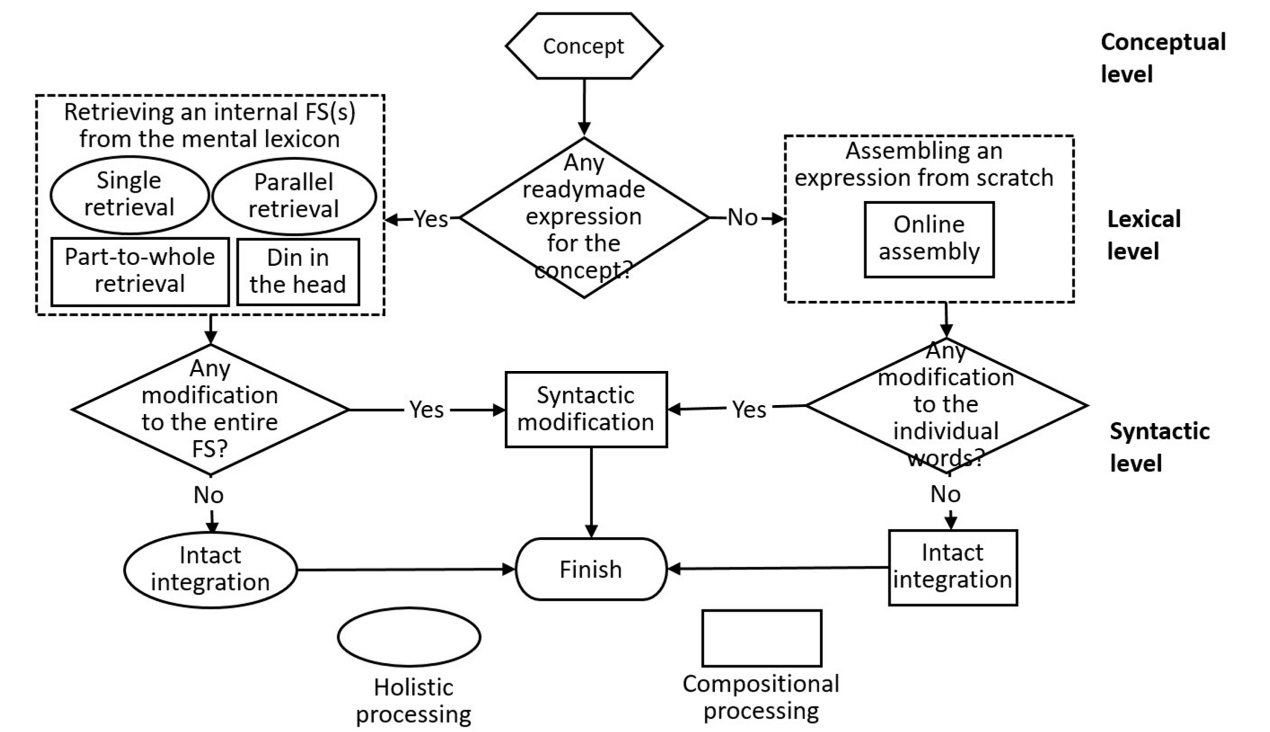


FIGURE 4
A model of L2 formulaic sequence production.

> Our teacher would note the phrases for us. I would mark them with the highlighter, so they would look like chunks. During morning reading, I would recite them. When reading passages, I myself would note down the phrases that seemed good and valuable. (HS3)

Contrastively, with lower awareness of FS memorization, lower-proficiency students tended to only passively memorize the FSs under teachers' requirement, as seen in the following interview excerpt:

> Our teacher would teach some key phrases. However, to recite or not, was up to us. For me who did not like reciting, I would glance at them over and over, cherishing every moment before the dictation. Then, shut the book and write. (LS2)

Briefly, higher-proficiency students were more like active accumulators of FS knowledge, while their lower-proficiency counterparts seemed like passive receivers in learning FSs. Therefore, the former would have richer FS storage, thus retrieving more internal FSs during writing. This finding is consistent with Chen (2019) result that the high-performance group used more recited collocations than the low-performance group. The author posited that the high-performance learners had developed the habit of reciting collocations as wholes, while the low-performance learners had much lower awareness of collocations. Both studies have underscored high-proficiency students' high awareness of FS memorization.

The reason why higher-proficiency students made more syntactic modifications to internal FSs may be explained as follows. First, those students probably used more formally flexible FSs such as phrasal verbs which were prone to occur with inflections and additions of modifying elements such as *have caused serious damage to*. By contrast, lower-proficiency students probably used a higher proportion of inflexible FSs, such as *of course*, *at the same time* and *the details are as follows*. It is also possible that they were more likely to store and retrieve fully lexically specified sequences (e.g., *from my perspective*), although these sequences can be seen as instantiations of flexible formulaic frames (e.g., *from…perspective*).

Second, higher-proficiency students might pay more attention to the contextual appropriateness of their FS use such as non-redundancy and grammaticality, so they would tailor the FSs more meticulously for the specific context by making more syntactic modifications. Lower-proficiency students, on the other hand, might sometimes lack consideration about the contextual appropriateness of FSs, thus embedding them directly in the text without proper modification. Similarly, Tavakoli and Uchihara (2020) noted that, more advanced learners could recycle FSs from task prompts competently by using them creatively in various forms, whereas lower-proficiency speakers used them repeatedly in the same form. Both studies have indicated the improved ability of adjusting the FSs to the specific contexts with increased L2 proficiency.

Third, higher-proficiency students might have better understandings of syntactic structures of the FSs, so they could manipulate them more flexibly and make more complex modifications. By contrast, lower-proficiency students might have inadequate understanding of the FS structures, so they could only use the FSs in a rigid way. For example:

> This expression (As far as I'm concerned) came out directly. I use it frequently. And there will be no worry about lexical error, because it is a fixed phrase. How nice! (LS2-2)

Acquisition experience:

> Our teacher marked out this expression and asked us to memorize it. Actually, I was wondering why adding these words up can mean 就我而言, but it does mean this.

As can be seen, due to inadequate understanding of the FS structure, the student would not fully parse the FS and make corresponding modifications to use it in a wider range of contexts.

The observed pattern that higher L2 proficiency led to more holistic processing on the lexical but not the syntactic level echoes the previous findings that L2 proficiency might not affect learners' FS processing in some aspects (Sonbul, 2015; Yeldham, 2022). However, this result only partially accords with the tendency discerned in Wolter and Yamashita (2018): the degree of holistic processing would increase with the development of L2 proficiency, as higher-proficiency learners showed more reliance on phrasal frequency than lower-proficiency learners.

Two possible reasons can account for such inconsistency. First, the judgment task with decontextualized test items in the previous study might only require minimal processing effort on the syntactic level, so the learners' performance seemed more reflective of the processing on the lexical level. Contrastively, the writing task in the current study demanded more processing effort on the syntactic level, i.e., embedding the FSs in the text, so it could be quite revealing of the learners' processing behaviors on this level. Second, the judgment task was administered in a controlled, time-pressured condition, so the learners might prioritize the processing on the lexical level, since it pertains mostly to meaning comprehension (Van Patten, 1990). Contrastively, the writing task in the current research was much more lenient in time. Thus, the learners, especially those of higher-proficiency, would be more consciously engaged in syntactic analysis of the FSs.

5.3 Effects of topic familiarity on learners' productive processing of FSs

The findings partially support the third hypothesis in that higher topic familiarity did not lead to more holistic processing on the lexical level, and it led to more holistic processing on the syntactic level. Specifically, in the familiar-topic writing, although the learners had retrieved more internal FSs in number, the proportion of internal FSs remained stable across the two tasks owing to the parallel increase of assembled FSs in number. Additionally, in the familiar-topic writing the learners made less modifications to the internal FSs.

The result that the learners had retrieved more internal FSs for the familiar topic can be explained from two perspectives: vocabulary readiness (abundant storage of FSs in the mental lexicon) and greater attention to form (more thorough search of the mental lexicon). First, the learners probably stored more FSs related to the familiar content domain. Previous research has shown that topic familiarity can bring about readiness in terms of content and vocabulary (Bui, 2014; Yang and Kim, 2020). Importantly, vocabulary readiness should be reflected in a rich storage of FSs semantically related to the familiar topic. Indeed, as recurrent FSs emerge in and thus depend on specific contexts (Mac Whinney, 2001), researchers have assumed that language users "master formulaic sequences associated with 'common' situations better than those occurring in unfamiliar situations" (Forsberg and Fant, 2010, p. 49).

Second, another benefit brought by topic familiarity seemed to be the greater amount of available attention to forms. According to Skehan's (1998) Limited Capacity Hypothesis, familiar tasks are less cognitively demanding, thus sparing more attentional resources for focus on form. Empirically, previous studies have found that familiar topics led to writing performance with higher lexical complexity (e.g., He and Shi, 2012; Yang and Kim, 2020). Indeed, we found that in the familiar-topic writing, students had made more attempts to upgrade their forms of expressions. For example, the student replaced *all kinds of* with *all sorts of* in his mind:

> To express 各种各样, I thought of the simple expression "all kinds of" initially. Then I felt "all sorts of" was more advanced. (HS5-Familiar)

Conceivably, in the familiar-topic writing, the learners would be more capable of attending to form and made more search for advanced expressions, even though the initial expressions were already workable. Thus, more stored FSs would be retrieved.

Nevertheless, the proportion of internal FSs remained stable across the two tasks, since the familiar-topic writing also involved more assembled FSs, which can be further explained by vocabulary readiness brought by topic familiarity as well. For the familiar topic, the learners might have stored more topic-related words, and also be relatively familiar with these words. Therefore, they seemed to have a higher chance of arriving at an acceptable expression through word-by-word assembly. Contrastively, in the unfamiliar-topic writing, without ready-made chunks, they had to frequently abandon the intended meaning or resort to non-formulaic language.

The fact that in the familiar-topic writing the learners made less modifications to the internal FSs might be explained as follows. For the familiar topic, the stored internal FSs might need little or no modification to be integrated into the current contexts, which might be highly consistent with their pervious contexts of use, as they probably had been used repeatedly in the same form to express certain familiar concepts, such as the use of the FS *it is universally acknowledged that* (HS5) as an introducer of familiar phenomena. By contrast, for the unfamiliar topic, the stored FSs may need more modifications to fit the unfamiliar contexts, which may differ considerably from their pervious contexts of use.

Nevertheless, concerning assembled FSs, a reverse situation was found: the learners made less modifications to them in the unfamiliar-topic writing. This is probably because the unfamiliar topic gave rise to more assembled FSs that denote abstract entities and thus do not need inflection in their usage. Specifically, these are the adjective-noun combinations for concepts in the economic field, such as *economic globalization*, *international cooperation*, *free trade*, and *the national economy*. Contrastively, in the familiar-topic writing, the assembled FSs for theme-related concepts tend to denote physical entities or behaviors, and thus occur with inflections. These include phrases describing students' daily life in explaining causes for stress (e.g., *social activities*, *complex issues*, *intellectual abilities*) and phrases about coping with stress (e.g., *playing games*, *playing sports*, *changing our attitude*). Briefly, concerning assembled FSs, the percentage of intact integration seemed to depend on how many of them denoted abstract concepts, which in turn was arguably determined by the abstractness of the topic, not necessarily the degree of familiarity. However, it might be assumed that the more abstract the topic, the less familiar the learners would be with it.

The observed pattern that higher topic familiarity led to more holistic processing on the syntactic level can help to explain the positive effect of topic familiarity on oral fluency detected in previous studies (Bui, 2014; Bui and Huang, 2018). Bui (2014) found that topic familiarity enabled learners to speak at a faster rate, with a longer mean length of run and fewer pauses. Bui and Huang (2018) showed that topic familiarity could reduce mid-clause pauses. As explained above, learners probably need not make much modification to internal FSs in the familiar contexts. Hence, their computational workloads would be reduced during the familiar-topic task, writing and speaking alike. Therefore, it is reasonable to assume that learners might make less online modifications to internal FSs in familiar-topic speaking tasks as well, thus promoting fluency.

5.4 Summary of major arguments

Regarding the first research question, this study proposed that the learners' FS processing types can be categorized on two levels: lexical and syntactic. On the lexical level, the learners engaged in the retrieval of internal FSs frequently, indicating that they have sizable storages of FSs. This lends support to previous findings that learners retain information about word co-occurrence. On the syntactic level, the learners engaged in syntactic modification of FSs to a considerable extent. This testifies to the syntactic flexibility of FSs detected in previous studies. Consequently, the learners had more holistic processing than compositional processing on the lexical level, but not on the syntactic level. In addition, a model of L2 FS production was proposed, which depicts the learners' FS processing types on different levels.

Regarding the second research question, this study proposed that, on the lexical level, higher-proficiency students engaged in the retrieval of internal FSs more frequently owing to their higher awareness of FS memorization and larger FS storage. On the syntactic level, they engaged in syntactic modification of FSs more frequently due to their use of more formally flexible FSs, greater attention to the contextual appropriateness of FSs and better understanding of FS structures. Consequently, higher L2 proficiency led to more holistic processing on the lexical but not the syntactic level.

Regarding the third research question, this study proposed that, on the lexical level, in the familiar-topic writing, learners engaged in the retrieval of internal FSs more frequently owing to the benefits of vocabulary readiness and greater attention to form brought by topic familiarity; however, they also engaged in more assembly of assembled FS, rendering the proportion of internal FSs unchanged. On the syntactic level, in the familiar-topic writing, learners engaged in less syntactic modification of internal FSs due to the presumably high consistency between the current and previous contexts of use for those FSs. Consequently, higher topic familiarity led to more holistic processing on the syntactic level but not the lexical level.

6 Conclusion

This study investigated L2 learners' processing of FSs in writing and the effects of L2 proficiency levels and topic familiarity on it. The learners' conscious processing (retrieval/assembly and integration into the text) of FSs was categorized on the lexical and syntactic levels, and these processing types were characterized as holistic processing or compositional processing. Results reveal that the learners retrieved FSs far more frequently than they assembled them, and made modification to about one thirds of the FSs. Furthermore, higher-proficiency students retrieved more internal FSs and made more modifications to them than their lower-proficiency counterparts; when writing on the familiar topic, the learners retrieved more internal FSs and had more intact integration of them.

Theoretically, this study bolsters our understanding of the cognitive processes involved in L2 FS production and contributes to the "holistic or compositional" debate on FS processing. Methodologically, it took full account of learner-internal FSs by training and inviting the participants to identify FSs. The training material can serve as a reference for future studies of learner-internal FSs.

Pedagogically, the two-layered categorization of FS processing types could help teachers better understand the causes of error in students' FS use, so they could prepare preventive measures in a more informed way. On the lexical level, students might combine words inappropriately, or retrieve an erroneous FS form due to memory lapse or incorrect fusion. On the syntactic level, they might make incorrect modifications to FSs or apply FSs too rigidly without contextual considerations. Besides, since FSs can be both holistic and compositional in production, teachers are advised to direct their students' attention to both the conventionality and flexibility of FSs. While emphasizing the importance of memorizing FSs as wholes, they can expose students to the contextualized uses of formally flexible FSs, and provide opportunities for them to use these FSs with different variations. Additionally, teachers can encourage students to memorize frequently-used, specified forms of some formulaic frames in order to reduce the computational workload during FS production.

Despite its contributions, this study suffered a number of limitations. The first was the small sample size. Future studies could recruit more participants from more different proficiency levels, or with different cognitive styles, to investigate the effects of learner-related factors on FS processing. Second, for the identification of learner-internal FSs, this study relied on the subjective judgments of the learners who might neglect some FSs due to lack of identification experience. Therefore, besides manual identification, further research could gather multiple compositions of the same learner and use concordance tools to extract the frequent sequences as a reference. In this way, FSs acquired more implicitly and thus easily neglected by the learners can be spotted. Third, this study used the method of VSR to investigate the learners' mental processes in FS production. Although VSRs were conducted immediately after the writing task, some details would inevitably be missed in the recall. Further research is suggested to examine the effectiveness of using VSR and think-aloud in combination. Finally, this study made a distinction between two processing levels to portray the FS processing types. However, the real cognitive processes in FS production might be more complex and defy such simple distinction. The two-layered categorization awaits further verification.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary material, further inquiries can be directed to the corresponding author.

Ethics statement

The studies involving humans were approved by Science and Technology Research Ethics Committee/Social Science Sub-Committees of Nanjing University. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

KF: Writing – original draft, Writing – review & editing, Conceptualization, Investigation. HW: Supervision, Writing – review & editing.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

Conflict of interest

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

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2024.1281926/full#supplementary-material

References

Abdi Tabari, M. (2022). Investigating the interactions between l2 writing processes and products under different task planning time conditions. *J. Sec. Lang. Writ.* 55, 100871–100816. doi: 10.1016/j.jslw.2022.100871

Ambridge, B., and Lieven, E. (2011). *Child language acquisition: Contrasting theoretical approaches*. Cambridge: Cambridge University Press.

Arnon, I., and Cohen Priva, U. (2014). Time and again: the changing effect of word and multiword frequency on phonetic duration for highly frequent sequences. *Mental Lexicon* 9, 377–400. doi: 10.1075/ml.9.3.01arn

Arnon, I., and Snider, N. (2010). More than words: frequency effects for multi-word phrases. *J. Mem. Lang.* 62, 67–82. doi: 10.1016/j.jml.2009.09.005

Barkema, H. (1994). "Determining the syntactic flexibility of idioms" in *Creating and using English language corpora*. eds. U. Fries, G. Tottie and P. Schneider (Amsterdam: Rodopi), 39–52.

Beck, S. D., and Weber, A. (2016). Bilingual and monolingual idiom processing is cut from the same cloth: the role of the L1 in literal and figurative meaning activation. *Front. Psychol.* 7:1350. doi: 10.3389/fpsyg.2016.01350

Biber, D., Johansson, S., Leech, G., Conrad, S., and Finegan, E. (1999). *Longman grammar of spoken and written English*. London: Pearson Education.

Bobrow, S. A., and Bell, S. M. (1973). On catching on to idiomatic expressions. *Mem. Cogn.* 1, 343–346. doi: 10.3758/BF03198118

Boers, F., Eyckmans, J., Kappel, K., Stengers, H., and Demecheleer, M. (2006). Formulaic sequences and perceived oral proficiency: putting a lexical approach to the test. *Lang. Teach. Res.* 10, 245–261. doi: 10.1191/1362168806lr195oa

Boers, F., and Lindstromberg, S. (2009). *Optimizing a lexical approach to instructed second language acquisition*. Basingstoke, UK: Palgrave Macmillan.

Bui, G. (2014). "Task readiness: theoretical framework and empirical evidence from topic familiarity, strategic planning, and proficiency levels" in *Processing perspectives on task performance*. ed. P. Skehan (Amsterdam: John Benjamins), 63–94.

Bui, G. (2021). Influence of learners' prior knowledge, L2 proficiency and pre-task planning on L2 lexical complexity. *Int. Rev. Appl. Linguist.* 59, 543–567. doi: 10.1515/iral-2018-0244/html

Bui, G., and Huang, Z. (2018). L2 fluency as influenced by content familiarity and planning: performance, measurement, and pedagogy. *Lang. Teach. Res.* 22, 94–114. doi: 10.1177/1362168816656650

Cacciari, C., and Corradini, P. (2015). Literal analysis and idiom retrieval in ambiguous idioms processing: a reading-time study. *J. Cogn. Psychol.* 27, 797–811. doi: 10.1080/20445911.2015.1049178

Cacciari, C., and Tabossi, P. (1988). The comprehension of idioms. *J. Mem. Lang.* 27, 668–683. doi: 10.1016/0749-596X(88)90014-9

Chen, W. (2019). Profiling collocations in EFL writing of Chinese tertiary learners. *RELC J.* 50, 53–70. doi: 10.1177/0033688217716507

Cieślicka, A. (2010). "Formulaic language in L2: Storage, retrieval and production of idioms by second language learners," in *Cognitive Processing in Second Language Acquisition*. Eds. M. Pütz and L. Sicola (Amsterdam & Philadelphia: John Benjamins), 149–168.

Cohen, A. D. (1998). *Strategies in learning and using a second language*. New York: Addison Wesley Longman.

Cordier, C. (2013). *The presence, nature and role of formulaic sequences in English advanced learners of French: A longitudinal study* Newcastle-upon-Tyne, UK: Newcastle University.

Cutting, J. C., and Bock, K. (1997). That's the way the cookie bounces: syntactic and semantic components of experimentally controlled idiom blends. *Mem. Cogn.* 25, 57–71. doi: 10.3758/bf03197285

Dempsey, N. P. (2010). Stimulated recall interviews in ethnography. *Qual. Sociol.* 33, 349–367. doi: 10.1007/s11133-010-9157-x

Divjak, D., and Caldwell-Harris, C. L. (2015). "Frequency and entrenchment" in *Handbook of cognitive linguistics*. eds. E. Dabrowska and D. Divjak (Berlin and New York: Mouton de Gruyter), 53–74.

Durrant, P. (2014). Corpus frequency and second language learners' knowledge of collocations: a meta-analysis. *Int. J. Corpus Linguist.* 19, 443–477. doi: 10.1075/ijcl.19.4.01dur

Durrant, P., and Schmitt, N. (2010). Adult learners' retention of collocations from exposure. *Second. Lang. Res.* 26, 163–188. doi: 10.1177/0267658309349431

Edmonds, A. (2014). Conventional Expressions: investigating pragmatics and processing. *Stud. Second. Lang. Acquis.* 36, 69–99. doi: 10.1017/S0272263113000557

Ellis, N. C., Simpson-Vlach, R., and Maynard, C. (2008). Formulaic language in native and second language speakers: psycholinguistics, corpus linguistics, and TESOL. *TESOL Q.* 42, 375–396. doi: 10.1002/j.1545-7249.2008.tb00137.x

Forsberg, F., and Fant, L. (2010). "Idiomatically speaking: effects of task variation on formulaic language in high proficient users of L2 French and Spanish" in *Perspectives on formulaic language: Acquisition and communication*. ed. D. Wood (New York: Continuum), 47–70.

Gass, S., and Mackey, A. (2017). *Stimulated recall methodology in applied linguistics and second language research*. New York: Routledge.

Gibbs, R. W. (1980). Spilling the beans on understanding and memory for idioms in conversation. *Mem. Cogn.* 8, 149–156. doi: 10.3758/BF03213418

Gibbs, R. W., and Nayak, N. P. (1989). Psycholinguistic studies on the syntactic behavior of idioms. *Cogn. Psychol.* 21, 100–138. doi: 10.1016/0010-0285(89)90004-2

Glucksberg, S. (1993). "Idiom meanings and allusional content" in *Idioms: Processing, structure, and interpretation*. eds. C. Cacciari and P. Tabossi (Hillsdale, NJ: Lawrence Erlbaum Associates), 3–26.

Grant, L. (2003). *A corpus-based investigation of idiomatic multi-word units* New Zealand: Victoria University of Wellington.

Hallin, A. E., and Van Lancker Sidtis, D. (2017). A closer look at formulaic language: prosodic characteristics of Swedish proverbs. *Appl. Linguis.* 38, 68–89. doi: 10.1093/applin/amu078

Hatami, S. (2015). Teaching formulaic sequences in the ESL classroom. *TESOL Q.* 6, 112–129. doi: 10.1002/tesj.143

He, L., and Shi, L. (2012). Topical knowledge and ESL writing. *Lang. Test.* 29, 443–464. doi: 10.1177/0265532212436659

Hoey, M. (2005). *Lexical priming: A new theory of words and language*. Oxford: Routledge.

Holsinger, E. (2013). Representing idioms: syntactic and contextual effects on idiom processing. *Lang. Speech* 56, 373–394. doi: 10.1177/0023830913484899

Huang, K. (2015). More does not mean better: frequency and accuracy analysis of lexical bundles in Chinese EFL learners' essay writing. *System* 53, 13–23. doi: 10.1016/j.system.2015.06.011

Hubers, F., Cucchiarini, C., and Strik, H. (2020). Second language learner intuitions of idiom properties: what do they tell us about l2 idiom knowledge and acquisition? *Lingua* 246, 102940–102914. doi: 10.1016/j.lingua.2020.102940

Jackendoff, R. (2002). *Foundations of language: Brain, meaning, grammar, evolution*. Oxford: Oxford University Press.

Jeong, H., and Jiang, N. (2019). Representation and processing of lexical bundles: evidence from word monitoring. *System* 80, 188–198. doi: 10.1016/j.system.2018.11.009

Jiang, N. (2000). Lexical representation and development in a second language. *Appl. Linguis.* 21, 47–77. doi: 10.1093/applin/21.1.47

Jiang, F. (2015). Nominal stance construction in L1 and L2 students' writing. *J. Engl. Acad. Purp.* 20, 90–102. doi: 10.1016/j.jeap.2015.07.002

Jiang, N., and Nekrasova, T. M. (2007). The processing of formulaic sequences by second language speakers. *Mod. Lang. J.* 91, 433–445. doi: 10.1111/j.1540-4781.2007.00589.x

Kessler, M., Ma, W., and Solheim, I. (2021). The effects of topic familiarity on text quality, complexity, accuracy, and fluency: a conceptual replication. *TESOL Q.* 56, 1163–1190. doi: 10.1002/tesq.3096

Kessler, R., Weber, A., and Friedrich, C. K. (2020). Activation of literal word meanings in idioms: evidence from eye–tracking and ERP experiments. *Lang. Speech* 64, 594–624. doi: 10.1177/0023830920943625

Kim, S. H., and Kim, J. H. (2012). Frequency effects in L2 multiword unit processing evidence from self-paced reading. *TESOL Q.* 46, 831–841. doi: 10.1002/tesq.66

Krashen, S. D. (1985). *Inquiries & insights: second language teaching: immersion & bilingual education, literacy.* Hayward, CA: Alemany Press.

Kyriacou, M., Conklin, K., and Thompson, D. (2020). Passivizability of idioms: has the wrong tree been barked up? *Lang. Speech* 63, 404–435. doi: 10.1177/0023830919847691

Langacker, R. W. (1987). *Foundations of cognitive grammar: Theoretical prerequisites*. Stanford, CA: Stanford University Press.

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

Lin, P. M. S. (2018). *The prosody of formulaic sequences: A Corpus and discourse approach*. London and New York: Bloomsbury Publishing.

Mac Whinney, B. (2001). "Emergentist approaches to language" in *Frequency and the emergence of linguistic structure*. eds. J. Bybee and P. Hopper (Amsterdam: John Benjamins), 449–469.

Mancuso, A., Elia, A., Laudanna, A., and Vietri, S. (2020). The role of syntactic variability and literal interpretation plausibility in idiom comprehension. *J. Psycholinguist. Res.* 49, 99–124. doi: 10.1007/s10936-019-09673-8

Millar, N. (2011). The processing of malformed formulaic language. *Appl. Linguis.* 32, 129–148. doi: 10.1093/applin/amq035

Moon, R. A. (1998). *Fixed expressions and idioms in English. A Corpus-based approach*. Oxford: Clarendon.

Myles, F., and Cordier, C. (2017). Formulaic sequence (FS) cannot be an umbrella term in SLA: focusing on psycholinguistic FSs and their identification. *Stud. Second. Lang. Acquis.* 39, 3–28. doi: 10.1017/S027226311600036X

Northbrook, J., and Conklin, K. (2019). Is what you put in what you get out? Textbook-derived lexical bundle processing in beginner English learners. *Appl. Linguis.* 40, 816–833. doi: 10.1093/applin/amy027

Öksüz, D., Brezina, V., and Rebuschat, P. (2020). Collocational processing in L1 and L2: the effects of word frequency, collocational frequency, and association. *Lang. Learn.* 71, 55–98. doi: 10.1111/lang.12427

Qi, Y., and Ding, Y. R. (2011). Use of formulaic sequences in monologues of Chinese EFL learners. *System* 39, 164–174. doi: 10.1016/j.system.2011.02.003

Roelofs, A. (1992). A spreading-activation theory of lemma retrieval in speaking. *Cognition* 42, 107–142. doi: 10.1016/0010-0277(92)90041-F

Sasaki, M. (2000). Toward an empirical model of EFL writing processes: an exploratory study. *J. Sec. Lang. Writ.* 9, 259–291. doi: 10.1016/S1060-3743(00)00028-X

Schmid, H.-J. (2007). "Entrenchment, salience and basic levels" in *The Oxford handbook of cognitive linguistics*. eds. D. Geeraerts and H. Cuyckens (Oxford: Oxford University Press), 117–138.

Schmitt, N., Grandage, S., and Adolphs, S. (2004). "Are corpus-derived recurrent clusters psycholinguistically valid?" in *Formulaic sequences: Acquisition, processing and use*. ed. N. Schmitt (Philadelphia, PA: John Benjamins), 127–151.

Shin, Y. (2019). Do native writers always have a head start over nonnative writers? The use of lexical bundles in college students' essays. *J. Engl. Acad. Purp.* 40, 1–14. doi: 10.1016/j.jeap.2019.04.004

Siyanova-Chanturia, A. (2015). On the 'holistic' nature of formulaic language. *Corpus Linguist. Linguist. Theory* 11, 285–301. doi: 10.1515/cllt-2014-0016

Siyanova-Chanturia, A., and Lin, P. (2017). Production of ambiguous idioms in English: a reading aloud study. *Int. J. Appl. Linguist.* 28, 58–70. doi: 10.1111/ijal.12183

Siyanova-Chanturia, A., and Martinez, R. (2015). The idiom principle revisited. *Appl. Linguis.* 36, 549–569. doi: 10.1093/applin/amt054

Skehan, P. (1998). *A cognitive approach to language learning*. Oxford: Oxford University Press.

Sonbul, S. (2015). Fatal mistake, awful mistake, or extreme mistake? Frequency effects on off-line/on-line collocational processing. *Biling. Lang. Congn.* 18, 419–437. doi: 10.1017/S1366728914000674

Spöttl, C., and McCarthy, M. (2004). "Comparing knowledge of formulaic sequences across L1, L2, L3, and L4" in *Formulaic sequences: Acquisition, processing and use*. ed. N. Schmitt (Philadelphia, PA: John Benjamins), 191–225.

Sprenger, S., Levelt, W. J. M., and Kempen, G. (2006). Lexical access during the production of idiomatic phrases. *J. Mem. Lang.* 54, 161–184. doi: 10.1016/j.jml.2005.11.001

Staples, S., Egbert, J., Biber, D., and McClair, A. (2013). Formulaic sequences and EAP writing development: lexical bundles in the TOEFL iBT writing section. *J. English Acad. Purp.* 12, 214–225. doi: 10.1016/j.jeap.2013.05.002

Stratman, J. F., and Hamp-Lyons, L. (1994). "Reactivity in concurrent think aloud protocols: issues for research" in *Speaking about writing: Reflections on research methodology*. ed. P. Smagorinsky (Thousand Oaks, CA: Sage), 89–112.

Swinney, D. A., and Cutler, A. (1979). The access and processing of idiomatic expressions. *J. Verbal Learn. Verbal Behav.* 18, 523–534. doi: 10.1016/S0022-5371(79)90284-6

Tavakoli, P., and Uchihara, T. (2020). To what extent are multiword sequences associated with oral fluency? *Lang. Learn.* 70, 506–547. doi: 10.1111/lang.12384

Tremblay, A., Derwing, B., Libben, G., and Westbury, C. (2011). Processing advantages of lexical bundles: evidence from self-paced and sentence recall task. *Lang. Learn.* 61, 569–613. doi: 10.1111/j.1467-9922.2010.00622.x

Underwood, G., Schmitt, N., and Galpin, A. (2004). "The eyes have it: an eye-movement study into the processing of formulaic sequences" in *Formulaic sequences: Acquisition, processing and use*. ed. N. Schmitt (Philadelphia, PA: John Benjamins), 153–172.

van Ginkel, W., and Dijkstra, T. (2019). The tug of war between an idiom's figurative and literal meanings: evidence from native and bilingual speakers. *Biling. Lang. Congn.* 23, 131–147. doi: 10.1017/S1366728918001219

Van Patten, B. (1990). Attending to content and form in the input: an experiment in consciousness. *Stud. Second. Lang. Acquis.* 12, 287–301. doi: 10.1017/S0272263100009177

Vincent, B. (2013). Investigating academic phraseology through combinations of very frequent words: a methodological exploration. *J. Engl. Acad. Purp.* 12, 44–56. doi: 10.1016/j.jeap.2012.11.007

Webb, S., Newton, J., and Chang, A. (2013). Incidental learning of collocation. *Lang. Learn.* 63, 91–120. doi: 10.1111/j.1467-9922.2012.00729.x

Wolter, B., and Gyllstad, H. (2013). Frequency of input and L2 collocational processing: a comparison of congruent and incongruent collocations. *Stud. Second. Lang. Acquis.* 35, 451–482. doi: 10.1017/S0272263113000107

Wolter, B., and Yamashita, J. (2018). Word frequency, collocational frequency, L1 congruency, and proficiency in L2 collocational processing: what accounts for L2 performance? *Stud. Second. Lang. Acquis.* 40, 395–416. doi: 10.1017/S0272263117000237

Wray, A. (2002). *Formulaic language and the lexicon*. Cambridge: Cambridge University Press.

Wray, A. (2008). *Formulaic language: Pushing the boundaries*. Oxford: Oxford University Press.

Wulff, S. (2008). *Rethinking Idiomaticity: A usage-based approach*. London and New York, NY: Continuum.

Wulff, S. (2018). "Acquisition of formulaic language from a usage-based perspective" in *Understanding formulaic language: A second language acquisition perspective*. eds. A. Siyanova-Chanturia and A. Pellicer-Sanchez (New York: Routledge), 19–37.

Xu, F. (2010). Retrieving patterns of lexical sequences by English majors in L2 timed writing. *Foreign Languages Teaching* 1, 22–26. doi: 10.13458/j.cnki.flatt.000289

Xu, F., and Ding, Y. R. (2010). Lexical-problem-solving strategies in L2 timed writing. *Foreign Languages China* 7, 54–62+111. doi: 10.13564/j.cnki.issn.1672-9382.2010.02.007

Yang, W., and Kim, Y. (2020). The effect of topic familiarity on the complexity, accuracy, and fluency of second language writing. *Appl. Linguist. Rev.* 11, 79–108. doi: 10.1515/applirev-2017-0017

Yeldham, M. (2022). Second language English listeners' relative processing of coherence-based and frequency-based formulas: a corpus-based study. *Appl. Linguist. Rev.* 13, 287–317. doi: 10.1515/applirev-2018-0093

Yoon, H. (2016). Association strength of verb-noun combinations in experienced NS and less experienced NNS writing: longitudinal and cross-sectional findings. *J. Sec. Lang. Writ.* 34, 42–57. doi: 10.1016/j.jslw.2016.11.001

Yu, Y. (2022). The role of psycholinguistics for language learning in teaching based on formulaic sequence use and oral fluency. *Front. Psychol.* 13:1012225. doi: 10.3389/fpsyg.2022.1012225

Yuan, H., and Xu, J. (2016). The psychological reality of automatic processing of formulaic sequences in L2 writing. *Foreign Language Edu* 37, 72–76. doi: 10.16362/j.cnki.cn61-1023/h.2016.01.016

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 24 April 2024
DOI 10.3389/fpsyg.2024.1305134

OPEN ACCESS

EDITED BY
Çise Çavuşoğlu,
Independent Researcher, Nicosia, Cyprus

REVIEWED BY
Andrzej Porzuczek,
University of Silesia in Katowice, Poland
Arkadiusz Rojczyk,
University of Silesia in Katowice, Poland

*CORRESPONDENCE
Malgorzata Kul
✉ kgosia@amu.edu.pl

RECEIVED 30 September 2023
ACCEPTED 01 March 2024
PUBLISHED 24 April 2024

CITATION
Kul M (2024) Perception of reduced forms in English by non-native users of English. *Front. Psychol.* 15:1305134.
doi: 10.3389/fpsyg.2024.1305134

Perception of reduced forms in English by non-native users of English

Malgorzata Kul*

Department of Contemporary English and Multilingualism, Adam Mickiewicz University, Poznań, Poland

The article reports the results of a study on the perception of reduced forms by non-native users of English. It tests three hypotheses: (i) reduced forms with context are recognized more accurately and faster than reduced forms without context; (ii) gradient reduction is perceived less robustly than the categorical one; and (iii) subjects with musical background perceive reduced forms better than those without. An E-Prime study on 102 Polish learners of English was implemented, comparing participants' accuracy and reaction times with a control group of 14 native speakers. The study was corpus-based and used 287 reduced forms from a corpus of Lancashire. The results indicate that (i) lexical context and phone density significantly affect perception, (ii) the category of reduction process (gradient or categorical) is irrelevant, and (iii) musical background only partially impacts non-native perception.

KEYWORDS

reduced forms, speech perception, E-prime, English as a second language (ESL), corpus-based study, Polish learners

1 Introduction

Due to a low degree of formality, attention (Labov, 1994), specific audience design (Bell, 1984, 2001), and high speech rate, casual speech abounds in reduction processes affecting both vowels and consonants. This, in turn, results in reduced forms (Shockey, 2003; Johnson, 2004). For instance, the phrase *I do not know /aɪ doʊnt noʊ/* assumes the reduced form / *dənoʊ/*.

Reduced forms, the topic of the study, are phonetic and phonological deviations from citation forms: "rapid speech—different word forms can emerge from rapid speech when compared with slow speech. For example, *perhaps* in clearly articulated slow speech becomes *praps* in rapid speech" (Bussmann, 1996, p. 396). According to Shockey (2003), "there are some phonological differences from citation forms […] I call these differences reductions" (Shockey, 2003, p. 1–2). Hanique et al. (2013) defined reduced forms as follows: "in conversational speech […] segments may be very short, altered […] or even completely absent" (Hanique et al., 2013, p. 1,644). One may understand reduced forms as a result of the operation of an array of phonological processes, either vowel reduction (i.e., centralization of vowel's formants and shortening of its duration, Lindblom, 1963) or consonant elision/assimilation/epenthesis, which occurs within and across a word boundary.

Processes occurring within words are relatively straightforward to explain as they occur within one lexical unit and have a relatively narrow domain. Cross-word phonological processes and ensuing reduction have been explored within the paradigm of the Production Planning Hypothesis (PPH, Wagner, 2012). Most theories (e.g., the variationist approach) seem to assume that all contextual information is available and retrieved simultaneously when a process transcends the boundaries of the word (Lange and Laganaro, 2014; Tamminga et al., 2016; Tamminga, 2018). PPH stipulates a more random way in which planning unfolds, departing from the linear assumption and leaning toward the view that only one word may be planned ahead at a time. PPH proposes that, apart from phonetic factors, phonological processes can be accounted for with a set of variables linked to production planning (called planning proxies), such as time pressure, speech rate, and degree of phonological/semantic/syntactic complexity of the following material. Other variables are probabilistic effects (Tanner et al., 2017; Kilbourn-Ceron et al., 2020) or duration of pause (Tanner et al., 2017).

The degree to which pronunciation of words may vary in colloquial speech due to phonological processes is impressive: Johnson (2004) reports that in spontaneously produced American English, over 60% of words deviate from their citation forms in at least one segment, whereas 6% of words miss at least one syllable in comparison with their full forms. For this reason, reduced forms may be potentially challenging for non-native speakers of English. Cruttenden (2008) observes that a second language is often learned on the basis of words in isolation and encourages EFL learners to familiarize themselves with assimilatory tendencies and weak forms. In a similar vein, Ernestus and Warner (2011) quote the form *yeshey* as a heavily reduced form of *yesterday*, stressing that reduced forms cannot be looked up in a dictionary by learners of English, nor can they be explained by native speakers who are usually not aware of reduction processes. Shockey (2003) points to a lack of significant contact with reduced forms when learners of a second language are taught by non-native speakers.

Apart from these problems, reduced forms related to consonants are governed by language-specific mechanisms, which add to difficulties in their production and perception. For instance, English and Polish (Dunaj, 1985, 2006; Madejowa, 1987, 1993; Madelska, 2005; Orzechowska, 2019; Zydorowicz, 2019) frequently exhibit the process of consonant cluster reduction (CCR), while in Greek, parallel CCR processes are rare. Shockey and Bond (2012) report that Greek learners of English find recognition of consonant cluster reduction challenging as they are not exposed to this particular variability in their L1. Polish learners of English, on the other hand, encounter fewer difficulties in recognizing English CCR processes. In a similar vein, a study by Shockey and Ćavar (2013) makes a number of phonotactics-related predictions, including those of Spanish, Latvian, and Greek learners of English.

1.1 Previous studies

Although the perception of reduced forms by native users of a language has been previously studied (e.g., Dilley and Pitt, 2010; Warner et al., 2012; Zimmerer and Reetz, 2014) and accounted for with usage-based and exemplar theories (e.g., Bybee, 2013), insights into the perception of casual speech by non-native speakers are infrequent in comparison. The vast majority of perception studies in language acquisition of casual speech, however, investigate vowel reduction, e.g., Smiljanic and Bradlow (2011), Sumner (2011), Van Dommelen and Hazan (2012), Ernestus et al. (2017), Brand and Ernestus (2018), and Morano et al. (2019). Thus, the area of consonant reduction appears to be underrepresented in research on reduced forms, with the notable exceptions of studies by Pearman (2004) and Shockey and Bond (2012).

The study by Pearman (2004) investigated the perception of casual English in L2 by Catalan learners, linking perception to the level of proficiency in English. Shockey and Bond (2012) tested the correct identification of reduction processes on Polish and Greek learners of English and concluded that the Polish learners outperformed the Greek ones. In particular, Pearman (2004) tested an array of reduction processes such as palatalization, place/manner assimilation, vowel/consonant weakening, and deletion. She gated the sentence *Is your friend the one that cannot go to bed by ten?* in 80 ms steps and presented it to 24 learners of English, both at the beginner and advanced levels of proficiency, and to a control group of 12 native speakers. She found that the outcomes of the advanced experimental group were consistently lower in comparison with the natives, and so were the scores of the beginner groups compared to the advanced one. Pearman (2004) also analyzed the level of confusion and concluded that with an increase in L2 experience, learners were perceptually readjusting in the direction of L2 phonological processing.

Similarly, Shockey and Bond (2012) used the gating technique in 50 ms gates on Greek and Polish learners of English. These groups were matched in terms of proficiency; thus, instead of proficiency level, language typology became a factor. To the authors' surprise, Polish learners of English scored nearly as high as native speakers and far surpassed Greek learners in recognition of reduced forms. Shockey and Bond (2012) concluded that "[a] possible reason is that Poles have similar syllabic patterns in their own language, while Greek syllable structure is entirely different and, on the whole, simpler" (Shockey and Bond, 2012, p. 208). Shockey and Ćavar (2013) explained this finding further with language-specific differences, claiming that Greek learners of English are not accustomed to recognizing reduced forms affecting, e.g., consonant clusters, as Greek has few of them.

In discussing previous research on the perception of reduced forms by non-native listeners, a number of distinct blind spots can be identified. First, much less is known about the mechanisms governing perception of reduction processes affecting consonants than vowel reduction. Second, the two studies reported above use rather scarce input to test the perception of second language learners, i.e., one read sentence that contains a small selection of reduced forms and was produced by only one speaker: *Is your friend the one that cannot go to bed by ten?* [ɪ**z** **j**ə frend ðə wʌn ðə**t** **k**æn**t** **g**oʊ **t**ə bed baɪ ten] (Pearman, 2004) and *So it was quite good fun, actually, at the wedding, though…* [səʊ ɪ**t** **w**əs kwaɪ**t** **g**ʊ**d** **f**ʌn ˈækʃəli ə**t** **ð**ə ˈwedɪ**ŋ** **ð**əʊ] (Shockey and Bond, 2012). Third, the relatively low number of participants in the two studies, 24 in Pearman (2004) and 31 in Shockey and Bond (2012), prevents making generalizations about the perception of casual speech by learners and calls for more research. Fourth, the two studies lack the variability of tokens and speakers used in their experiments. In the spirit of corpus phonology (CP), the employment of a speech corpus seems a more suitable choice and would allow the drawing of more reliable generalizations than one sentence produced by one speaker. CP may be explained as "a novel methodological approach in phonology, denoting the use of purpose-built phonological corpora for studying speakers' and listener' knowledge and use of the sound system of their native language(s), the laws underlying such sound systems, L1 and L2 acquisition" (Gut and Voormann, 2014, p. 13).

To address these urgent neglects, a study was designed that seeks to fill the gap of using scarce input to investigate a wider array of consonant reduction processes and test a large number of tokens from a speech corpus on a significant number of subjects. The study reports the results on the perception of reduced forms in English by Polish learners of English and considers exclusively reduction processes affecting consonants: /t, d, h/ deletion, fricativization, assimilation of the place of articulation, and Yod coalescence. In particular, the study verifies the accuracy (referred to as Acc) and speed (reaction time, RT) of perception of reduced forms and makes a contribution to the area of perception of English casual speech by non-native users via the employment of a speech corpus.

Previous studies on native perception have typically addressed vowel reduction from the following angles: segmental context (Mitterer and Ernestus, 2006; Mitterer et al., 2008; Zimmerer and Reetz, 2014), word context (e.g., Van De Ven et al., 2011), word probability (van de Ven et al., 2012), speech rate (Dilley and Pitt, 2010), phonotactics (Spinelli and Gros-Balthazard, 2007), and syntax (e.g., Viebahn et al., 2015). As for the perception of vowel reduction by non-native listeners, the variables were as follows: high proficiency learners (Nouveau, 2012; ten Bosch et al., 2016; Wong et al., 2017), phonotactics (Shockey and Ćavar, 2013; Ernestus et al., 2017), speech styles (Smiljanic and Bradlow, 2011), voice onset time (Sumner, 2011), vowel formants and duration (Van Dommelen and Hazan, 2012), frequency of occurrence and exposure to a word (Brand and Ernestus, 2018), and word exemplars effects (Morano et al., 2019). Regarding the non-native perception of consonantal reduction, the following factors have been considered so far: the proficiency level of learners (Pearman, 2004), linguistic typological differences (Shockey and Bond, 2012; Shockey and Ćavar, 2013), and a stay abroad in an English-speaking country (Shockey and Bond, 2012).

Considering that the factors listed above were investigated mostly in native speech and concerned vowel reduction, the present study aims to advance our understanding of the non-native perception of consonantal reduction by conducting a corpus-based analysis. Due to the paucity of studies on learners' perception of reduced forms relative to native speakers' perception, the study analyzes non-native listeners' performance with native speakers as a control group (Pearman, 2004; Shockey and Bond, 2012). Instead of replicating the well-documented variables governing perception, the study explores the effects of semantic context (Van De Ven et al., 2011), the effects of a process type whose impact is known from production studies (e.g., Holst and Nolan, 1995) but has not yet been tested in the perception of reduced forms. The study also tests the effect of music education, which might influence the perception of casual speech as it does aid the production of segments (e.g., Magne et al., 2006; Milovanov et al., 2010; Salcedo, 2010).

1.2 The effects of semantic context

A number of production studies found that learners often rely on the meaning of the semantic context (e.g., Van De Ven et al., 2011). The present study extends this claim to the perception of reduced forms and includes the effects of lexical context, understood as the presence or absence of words preceding and following the reduced form in question. It aims to establish whether the presence of context facilitates the perception of reduced forms, as opposed to the absence of context (e.g., *they should have do**n**e **b**etter* vs. *do**n**e **b**etter,* place assimilation of /n/ to /m/ in the vicinity of a bilabial sound), which in turn derives from claims that L2 learners require a lot more acoustic signal such as vowel formants and formant transitions for consonants and/or a greater portion of the word than native speakers. Learners often need to hear the beginning of the next word before correctly identifying an item; in comparison, native speakers usually recognize the word before its offset (Nooteboom and Truin, 1980; Koster, 1987).

1.3 The effects of process type

A wide range of phonological processes is responsible for the difference between citation and reduced forms. This variety is language-specific and largely depends on the phonemic inventory of a language, word stress rules, and phonotactic constraints. Processes themselves operate in three ways: they can add a sound (*ham(p)ster*), delete it (*min**d** the gap*), or assimilate two or more sounds (*coul**d** **b**e*). Consequently, a process can belong to one of three major process groups: insertion, elision, and assimilation. Since they exert different phonetic effects, phonological accounts distinguish between categorical (e.g., Clements, 1985; McCarthy, 1988) and gradient types of processes (Browman and Goldstein, 1990; Barry, 1992). Assimilatory processes illustrate the gradient type as a change from sound A to sound B, which may involve intermediate stages, be incomplete, and leave a phonetic trace. Previous studies report that, e.g., an alveolar stop assimilating to a following labial or velar does show acoustic (Gow, 2003) or articulatory (Ohala, 1990; Nolan, 1992; Zsiga, 1995; Ellis and Hardcastle, 2002) traces of both alveolar and labial/velar place of articulation. We also know that extreme assimilation (i.e., complete blending of two places of articulation) results in considerable processing difficulties (Holst and Nolan, 1995). Categorical type of processes such as elision, on the other hand, neither leave a trace nor change one sound into another. Instead, they involve a complete realization. Studies on L1 production (Wright and Kerswill, 1989; Ellis and Hardcastle, 2002) and L1 perception (Hanique et al., 2013) attest to different effects of gradient and categorical processes, respectively. In this connection, the current study aims to explore the effects of process type on non-native perception. The idea that gradient processes have a different effect on learners than the categorical one seems worth pursuing.

1.4 The effects of musical background

The study investigates the effects of musical education, formal or informal (including singing), which may affect the perception of reduced forms. The claim that musical aptitude and linguistic skills are interconnected is well evidenced (e.g., Mithen, 2005; Jackendoff, 2009; Harvey, 2017), especially in the area of rhythm (Patel, 2003; Patel and Daniele, 2003; Mora and Gant, 2016). It has also been assumed that learners with musical talent and training achieve better results in pronunciation than learners without a musical background (e.g., Magne et al., 2006; Pastuszek-Lipińska, 2009; Milovanov et al., 2010; Salcedo, 2010). In the light of emerging evidence, music aptitude in SLA positively affects the production of certain aspects of English-connected speech:

rhythm, elision, assimilation, and linking (Milovanov, 2009; Besson et al., 2011; Gordon et al., 2011; Balčytytė-Kurtinienė, 2018). Casual speech mostly consists of vowel reduction, specific alternation of stressed and unstressed syllables, and an abundance of weak forms and weak syllables; the study by Gordon et al. (2011) examined the perception of weak and strong syllables in songs, pointing to the role of songs in increased attention of listeners when the beat was aligned with a strong syllable. In the study by Besson et al. (2011), musicians and non-musicians were compared in their perception performance of pitch, vowel duration, and metric processing in casual speech by training transfer and demonstrated overall facilitation for musicians. As far as rhythm and duration of vowels are concerned, the study by Milovanov (2009) proved that learners with greater musical aptitude had better scores in recognition and discrimination tasks than learners with no such skills. Following these suggestions, the current study tests this claim in the perception of consonantal reduction in casual speech. The choice of this particular effect was also motivated by the willingness to incorporate one extralinguistic factor in the study, in addition to the two linguistic ones (i.e., the effects of lexical context and process type), since previous studies on native perception of reduced forms and their variants examine speech rate as an extralinguistic factor (Dilley and Pitt, 2010).

To sum up, the present study analyzes the perception of reduced forms by learners of English to test the effects of (i) lexical context, (ii) two types of reduced forms (categorical and gradient), and (iii) musical background. The second aim determines the choice of phonological processes as they need to represent both gradient and categorical types: /t, d, h/ deletion, fricativization, assimilation of the place of articulation, and Yod coalescence. The study also aims to compare the learner's results to those of a control group of native speakers, and it was designed to use authentic, spoken English from a speech corpus in an SLA study.

1.5 The hypotheses

Pursuing the above aims, the study verifies three hypotheses:

Hypothesis 1 concerns the effects of lexical context on the perception of reduced forms and predicts that reduced forms with context are recognized more accurately and faster than reduced forms without context by learners, as reported in previous works (Nooteboom and Truin, 1980; Koster, 1987; van de Ven et al., 2012).

Hypothesis 2 addresses the effects of a process type. It is not unreasonable to expect that L2 listeners have a sharper perception of words with a missing sound (categorical processes) than for the potentially confusing effects of gradient processes such as place assimilation (e.g., change from /d/ to /g/), Yod coalescence (replacing /t, d, s, z/ in the vicinity of /j/ with /ʃ, ʒ, tʃ, dʒ /), or fricativization (/h/−like sound instead of a stop), which might cause perceptual difficulties. Thus, hypothesis 2 assumes that gradient reduction (fricativization, place assimilation, and Yod coalescence) is perceived slower and less accurately than the categorical one (/t, d, h/ deletion; Holst and Nolan, 1995; Hanique et al., 2013).

Hypothesis 3 extends the role of musical background from production and perception of individual sounds and suprasegmentals to casual speech (e.g., Besson et al., 2011). It stipulates that learners with musical backgrounds perceive reduced forms more accurately and in a shorter time than those without them, testing the effect of music skills or education on non-native perception.

2 Materials and methods

2.1 Subjects

One hundred and two adult Polish learners of English in the full- and part-time English and Russian-English programs at Adam Mickiewicz University participated in the study. Ninety-nine subjects ranged in age from 18 to 24 years, with a mean age of 21 years. The remaining three participants were aged 17, 37, and 40 years. All of them were native speakers of Polish at an English proficiency level, which can be described as advanced. In particular, the subjects were continuing their education in the second year of their postgraduate studies. Following their first year, they had to take an exam, placing them at advanced level described as B2+ (University of Adam Mickiewicz, n.d.). For this reason, the study cannot include proficiency level due to the highly even level of English among the subjects considered: they all were advanced students of English.

In addition, the participants had a similar background in phonetics since all of them had attended two obligatory courses in pronunciation of English: an EFL course in pronunciation and a course in phonetics and phonology. The former consists of drilling vowels, diphthongs, and consonants with the use of a multimedia program (Sawala et al., 2024). The latter is of a theoretical nature, introducing concepts such as phonemes, phonotactics, phonostylistics (including reduction processes), and phonological theories.

To tease apart the effects of context in the perception of reduced forms, the participants were assigned to one of two groups: 52 subjects listened to reduced forms without context (the NoContext group), whereas the remaining 50 participants constituted the Context group, which heard the same reduced forms as the NoContext group, but in context. A control group of 14 native English listeners served to demonstrate that the L2 learners were affected by processes of connected speech and not some other factors that might be present in the signal. Thus, 8 native listeners were presented with the stimuli in context, and 6 native speakers of English listened to the stimuli without lexical context. Table 1 presents their accents.

One may wonder whether accent variability, reported in Table 1, might exert an influence on the comparison between native and non-native listeners. With the exception of fricativization being reported mostly for the UK dialects (Lodge, 1984; Beal, 2010), all remaining reduction processes (h, t, d-deletion, assimilation of place, and Yod coalescence) occur in English regardless of the dialectal differences (Wells, 1982; Kortmann and Upton, 2008). For instance, h-dropping is a function of grammar (e.g., unstressed personal pronouns), and place assimilation is conditioned by the phonetic context and is not specific to a variety. As a textbook variable of sociolinguistic variation, /t, d/ deletion has been investigated in a number of varieties, such as American English in general (Neu, 1980; Guy, 1992; Guy and Boberg, 1997), in particular Philadelphia English (Guy, 1980), New York English (Labov et al., 1968), Appalachian English (Hazen, 2011), African American Vernacular English (Wolfram, 1969; Fasold, 1972), and British English (Tanner et al., 2017), specifically York English (Tagliamonte and Temple, 2005).

In addition, the study examines an extralinguistic factor related to the subject background, i.e., musical background by which the study understands the ability to play an instrument, either self-taught or obtained via formal music education and/or singing in a choir or a band. Musical background was established using a questionnaire, and

TABLE 1 Native speakers' places of residence.

Control group for the NoContext group		Control group for the Context group	
Initials	Place of residence (by 18)	Initials	Place of residence (by 18)
PN	UK (Manchester)	DL	UK (Solihull)
TA (a)	US (California, Oklahoma, Ohio)	SN	Canada (Alberta)
BF	UK (Leeds, Hull)	TdP	Australia (Western Australia)
TA (b)	Ireland	RH	Canada (British Columbia)
SB	UK (Norfolk, Essex)	CP	UK (South)
CW	UK (Oldham)	SH	UK (North Yorkshire)
		BN	US (California)
		RJ	US (Oregon)

results are reported as follows: 27 subjects from NoContext Group (52%) had musical background. However, only for 6 participants (22%) was the education formal. The duration of years spent in musical school ranged from 4 to 7 years and referred to the first degree (5.5. years on average). In Poland, there are three types of musical schools: first-degree and second-degree musical schools, which require 4 to 6 years of training, and third-degree (School of Music, 6 years) musical schools. In comparison with informal musical education (e.g., private lessons), the musical school offers an intensive program (courses at least twice a week), including theoretical modules such as rhythm, harmony, and composition. In the Context Group, 36% (18 participants) claimed to have some musical background. Within the group, 66% (10 subjects) received a formal musical education, both the first and second degrees. Duration of schooling ranged from 2 to 9 years, with an average of 5.4 years.

To sum this subsection up, the study tests the perception of reduced forms in English in the following experimental setting: the analysis is performed on two groups of subjects (learners and native users of English) and aims to tease apart the effects of lexical context (hence, four groups: NoContext leaners, NoContext native speakers, Context learners, and Context native speakers) as well as the effects of musical background in perception (for learners only, four groups: NoContext group musical background, NoContext group no musical background, Context group musical background, and Context group no musical background).

2.2 Speech material

The study was corpus-based and used reduced forms from the Phonologie de l'Anglais Contemporain corpus (PAC, Durand and Pukli, 2004). It contains recordings of 9 female speakers of Lancashire, collected between 2001 and 2002. PAC's structure is as follows: a list of words, a read passage, and formal and informal interviews. Both interviews were loosely structured and conducted in an informal setting, at informants' homes or workplaces. The formal interview was conducted by a French speaker of English, a stranger to the informants. The informal interview, on the other hand, was carried out by a native speaker of Lancashire who was either a relative or a friend (or a family friend) of the informants. Due to these differences, topics in the formal interview covered past events, memories from school, family situations, jobs, and travels abroad. The informal part concerned current topics such as housing problems, plans for Christmas, and gossip about common friends, relatives, and neighbors. For this study, the speech material comes from informal interviews to ensure fully casual speech. Although the choice of the English accent in this study may seem somewhat arbitrary, one particular argument speaks in favor of PAC: the PAC corpus of Lancashire was deemed appropriate since it has already been annotated and exploited with respect to reduced forms. Other corpora of spoken English are much less easy to mine as none of them are annotated beyond segments.

With regard to stimuli, particular care was taken to select simple words and phrases in which a reduction process occurred either within a word or across word boundaries. Lexical items used to test reduction were nearly all high-frequency items (following the British National Corpus, 2017, for frequency rankings, *cf.* Appendix 1), which excludes the possibility that the subject did not understand the word(s) they were supposed to identify in reduced forms.

The study's author manually cut the stimuli from the PAC corpus using Audacity software and exercised caution to include, e.g., the release stage of a stop in the signal. The stimuli, fed in the WAV format into an E-prime script, were high-quality recordings in the .wav format. In total, the study tested the perception of 287 stimuli, representing six tested processes: /t/ deletion, /d/ deletion, /h/ deletion, fricativization, place assimilation, and Yod coalescence. The NoContext group listened to 170 reduced forms without context (30 per each of six categories with the exception of fricativization, totaling 20 instances since most of the examples of fricativization in the corpus were preglottalized), whereas the Context Group heard 117 stimuli (20 per each process and 17 for fricativization) with context; the stimuli with lexical context were longer, and for this reason, their number was lower. It has to be clarified that context does not entail a full sentence, as sentences in casual speech are frequently long, complex, unfinished, or interrupted. Instead, a smaller portion of an intonation phrase within a breath group and defined by pauses were selected by hand, ranging from one to 15 following and preceding words. All stimuli were presented in random order for every subject. Table 2 exemplifies the stimuli; for a full list, see Appendices 1, 2.

2.3 The procedure

Participants completed the experiment in the Language and Communication Laboratory at a Polish state university in a single session. Before the study, they filled in a questionnaire that furnished metadata on the subjects and their language history. The questionnaire

TABLE 2 Examples of stimuli.

Process	Stimuli without context	Stimuli with context
/t/ deletion	Must be	They must be altered
/d/ deletion	Pounds	Nearly 2,000 pounds
/h/ deletion	Seeher	I still seeher occasionally
Fricativization	Obviously	Obviously I had a family to look after
Place assimilation	Ankle	Then I've sprained my ankle
Yod coalescence	Theseyears	But you have remembered it all theseyears

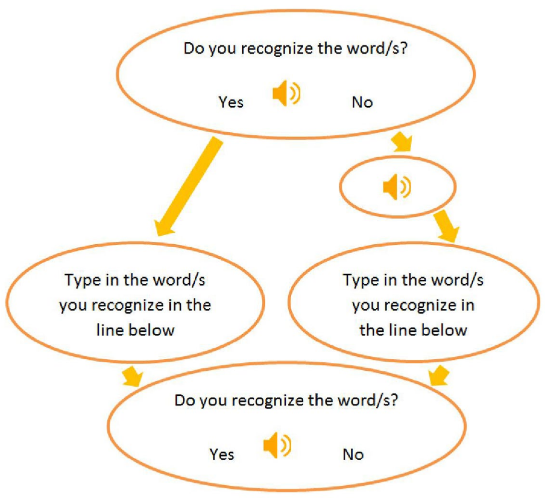


FIGURE 1
The E-prime task.

included sections about the subjects' musical backgrounds. Next, the subjects proceeded to perform the experiment in E-prime 2.0, equipped with Sennheiser headphones. The script in E-prime consisted of a trial session, a 2-min-long recording which introduced the subjects to the Lancashire dialect, speakers, and the task. The task contained no visual elements except for the icons of a loudspeaker to indicate a sound file. The instruction was the same for both groups, and the subjects were asked whether they recognized the word/s they heard (measuring reaction time) and to type in the words they recognized (to capture accuracy). Failure to provide a Yes answer indicated a lack of understanding of a stimulus and resulted in hearing the stimulus once more; each stimulus could be played twice. Only then did the next screen appear with a box to type. The task is visualized in Figure 1.

The subjects were not aware of the purpose of the study; they only knew that they were going to listen to a few samples of Lancashire. Participation was voluntary and took place before or after their regular courses.

The study analyzed a total of 14,690 answers provided by Polish learners of English, as well as 1956 answers from native English speakers. Although the E-prime script had automatically assigned points for correct answers, the study's author ran an independent, manual analysis of the obtained results to include answers with typos or orthographical errors. The analysis was binary: a point was assigned if the answer demonstrated that the subject had correctly understood the reduced form in question (even if the rest of the token was not understood correctly), and zero points if there was a failure to comprehend the meaning of the reduced form. That involved cases where L2 learners typed correctly a word or two words affected by a process and the preceding and upcoming words were not understood.

The measurements of RTs were triggered when the subjects pressed the Yes/No button. In particular, RTs were measured for both No and Yes answers by default. In the course of analysis, the results from the No column in the spreadsheet generated by the script were not taken into account. Following this step, I also performed a manual verification of the answers in the Yes column, where the subjects were asked to write what they had heard. The study's participants were not specifically instructed to provide the IPA type of transcription of the words they heard and supposedly recognized. Some subjects attempted some sort of impressionistic transcription, and others used an orthographic one. It has turned out that certain subjects prescriptively corrected a reduced form, whereas other participants managed to capture the effects of a process. For these reasons, I had to read every single typed response and decide whether the words were recognized correctly as well as exclude the" mondegreens" mishearings. For instance, *makes you* was frequently written down as *makes sure*. The final step of that manual verification was another column in the spreadsheet annotating if this response (and the RT) could be considered in the analysis. Categorical/gradient reduction was measured in the same way as overall word recognition, by hand, based on the typed response with filtered deletion or non-deleting processes.

One-way ANOVA was run to assess the differences between accuracy and reaction time between various groups, i.e., Context and NoContext groups, native speakers and learners, gradient and categorical type of processes, and musical backgrounds. Next, the effect size is reported as Cohen's d for equal groups and slightly modified for comparison between native and non-native listeners due to the difference in sample size with Hedges' g. Two-way ANOVA (without replication) was run to assess the difference between subjects with and without music education since there were two independent variables, i.e., musical background (Hypothesis 2) and lexical context (Hypothesis 1).

3 Results

3.1 Results for hypothesis 1

The first hypothesis predicted that the presence of lexical context boosts the perception of reduced forms. Figure 2 presents the results.

Regarding accuracy, although a difference between learner's Context and NoContext groups was observed, $F(1,100) = 52.76$, $p < 0.05$ (NoContext: M = 53.25, SD = 10.56, $N = 52$; Context group:

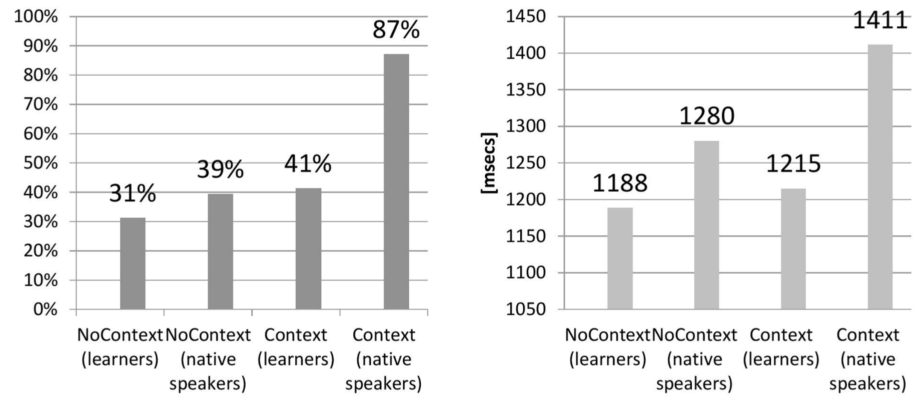


FIGURE 2
Comparison of Acc and RT between groups.

M = 48.42, SD = 12.01, N = 50), the value of Cohen's effect size for the differences between learners' groups (r = *0.21*, d = 0.43) suggested small practical significance between the context and no-context conditions (31% and 41%, respectively). Concerning the comparison between learners (NoContext: M = 53.25, SD = 10.56, N = 52, Context: M = 48.24, SD: 12.01, N = 50) and native speakers (NoContext: M = 67, SD = 12.02, N = 6; Context: M = 102, SD = 7.97, N = 8), there was a difference for the NoContext Group, $F(1,75) = 94.58$, $p < 0.05$, as well for the Context Group, $F(1,75) = 1720.88$, $p < 0.05$. Again, contrary to the results of a mere comparison, it turns out that the effect size for the NoContext learners vs. NoContext native speakers groups was weakly significant (r = 0.34, d = −1.18, g = *−1.26*). It was highly significant, d = −5.26, r = 0.88, g = −4.56, only for the Context groups of learners and native speakers. In the analysis, the strength of the association (r) was calculated using d for unequal groups.

The variation of the effect size in the Context vs. NoContext conditions merits further analysis, which reports exact differences between the groups of learners and native speakers. To this end, Tukey's Honest Significant Difference test (HSD) was run as a post-hoc test. A Tukey's *post hoc* test revealed that with an error rate of 0.05, there was no statistically significant difference between the learners' Context and NoContext groups as their confidence interval contained zero (difference points = −4.83, p = 0.14), unlike for native speakers (difference points = 35, p = 0.00). The post-hoc test also demonstrated that the improvement in the performance was significantly lower for the learner vs. native speakers groups in the NoContext condition (difference points = 13.75, p = 0.03) compared to the difference between learners and native speakers in the Context condition (difference points = 53.58, p = 0.00). An interesting observation is that the difference between learners and native speakers in the NoContext condition was statistically significant. This implies that native users of English outperformed the learners even if the lexical context was missing.

Figure 2 also suggests between-groups differences for reaction time, which were then investigated with one-way ANOVA. The compared groups of learners differ significantly, $F(1,100) = 420.27$, $p < 0.05$, allowing us to observe that the Context Group (M = 1359.31 ms, SD = 786.17, N = 50) recognized reduced forms faster than the NoContext Group (M = 1466.27 ms, SD = 605.21, N = 52). In comparison with native speakers, tokens without context were recognized slower by learners than by native speakers, $F(1,75) = 314.26$, $p < 0.05$ (learners: M = 1466.27 ms, SD = 605.21, N = 52; native speakers, M = 1284.24, SD = 179.36, N = 6). The same tendency was observed for the Context Group: $F(1,75) = 24.95$, $p < 0.05$ (learners: M = 1359.31 ms, SD = 786.17, N = 50; native speakers, M = 1411.28, SD = 469.29, N = 8). The effect size between two non-native speakers groups was not significant: r = −0.07, d = −0.14. With regard to the comparison of NoContext and Context between native speakers and learners, the effect size failed to reach any significance: r = 0.13, d = 0.41, g = 0.31 and r = 0.03, d = −0.08, g = −0.07, respectively.

Having applied a post-hoc test, Figure 2 can now be correctly interpreted. Thus, the study verifies Hypothesis 1 as follows: learners tested in the Context group recognized reduced forms faster but not more accurately than learners in the NoContext group. This outcome strongly suggests that the inclusion of semantic context did not help learners correctly identify reduced forms. In contrast, within the group of native speakers, semantic context did not shorten their reaction time but significantly improved their performance in comparison with the no-context group. Comparing the performance of native vs. non-native groups, native speakers have considerably outperformed learners in the perception of reduced forms, both with context and without context. Overall, hypothesis 1 was positively verified for native speakers of English and negatively for learners.

Following statistical analysis in terms of Cohen's effect size and *post hoc* Tukey HSD test, the outcomes obtained for hypothesis 1 reveal a significant difference between L1 and L2 users, which merits further investigation. In an attempt to cast light on this difference, the study explores two alternative suggestions that might account for the poor performance of Polish learners: the length of reduced form (calculated as phone density) and the influence of L1. These two possible explanations are discussed in turn.

Hypothesis 1, pertaining to the effects of lexical context, brought an interesting result for the learner group: even though the presence of context did not significantly aid accuracy in comparison with native users of English, the reaction time of learners was shorter than that of native speakers (Figure 2). This raises the question about the role of cognitive load since the tokens varied greatly in length due to the presence or absence of context. Tokens without context were consequently considerably shorter than the ones with context; still, there were differences between the two as the tokens with context were extracted from the corpus along natural boundaries constituted by intonation contours as well as by semantic/syntactic units and as a result were not of equal length. Given the limitations of working memory and increased cognitive load associated with longer phrases for learners, it can be expected that the length of a token has an influence on reaction time and, perhaps, accuracy in perception. In this study, semantic context or its lack determined the length of a token. As a follow-up on Hypothesis 1, the study now attempts to establish the influence of token length on perception operationalized as phone density and analyzed across low-, mid-, and high-density groups. Phone density was calculated as the number of phonemes in the phrase or sentence that contained a reduced form. For the NoContext group, it denoted the total number of phonemes in the token. Tokens with context were significantly longer; consequently, the sum of phonemes preceding and following the reduced form (but excluding the form itself) constituted phone density for the Context Group. The rationale behind this relied on the assumption that low density triggers higher accuracy and shorter reaction time (and that high density hinders correct perception and reaction time) and consists of correlating accuracy with reaction time for correct answers with three groups: tokens of low, mid, and high phone density.

The numbers for density groups (high, mid, and low) represent the actual numbers of phones in the phrases from the stimuli. In selecting the stimuli for the study, I used an intonation phrase or a pause as a unit/boundary so that the context would make a discourse unit as well. The criteria for assigning the number of phones to a particular group were as follows: the groups should not overlap nor be too close, and they should make generalizations possible.

Table 3 presents density groups, whereas Figure 3 summarizes the results. The results represent the means of correct answers per each density group (low, mid, and high) as well as means of accuracy and reaction times across density groups for learners.

The results for Context Group (low-density: M = 25, SD = 15, N = 9; mid-density: M = 19, SD = 14, N = 10; high-density: M = 14, SD = 11, N = 10) display the tendency toward better recognition of reduced forms in shorter tokens; for NoContext (low-density: M = 2, SD = 1, N = 7; mid-density: M = 11, SD = 7, N = 6; high-density: M = 4, SD = 7, N = 7), mid-density seems to boost perception. The results are statistically significant (NoContext Group: $F(2,17) = 33.86$, $p < 0.05$; Context Group: $F(2,26) = 21.61$, $p < 0.05$).

Reaction time neatly dovetails with the number of phonemes in the tokens: the longer the stimuli, the more time it took for learners to process a reduced form [(NoContext Group: $F(2,17) = 909.62$, $p < 0.05$; Context Group: $F(2,26) = 3840.74$, $p < 0.05$). No Context: (low-density: M = 838, SD = 455, N = 7; mid-density: M = 913, SD = 479, N = 6; high-density: M = 1,481, SD = 1,797, N = 7) and Context (low-density: M = 845, SD = 220, N = 9; mid-density: M = 932, SD = 268, N = 10; high-density: M = 1,840, SD = 1,770, N = 10)].

The length of a reduced form had an impact on perception only in the Context group (Figure 3), and this has been trending in the expected direction: the longer the phrase, the more difficult identification was. For the NoContext group, on the other hand, the role of phone density is not so clear, as low and high density produced similarly low results. Perhaps the group without lexical context was insensitive to the length of the form, and the mild effect of mid-phone density is purely accidental as it does not quite follow the pattern evidenced for the Context group. Nevertheless, it seems that the phone density account only partly predicts the perception of reduced forms by learners of English. Overall, the analysis of phone density as a possible predictor of perception has confirmed the intuitive connection between phone density and reaction time in the learner's group: the shorter the token, the shorter the reaction time, regardless of the presence or absence of context. Regarding accuracy, the same relationship was observed for the context group: the longer the token, the worse the perception. The NoContext group, however, does not follow this tendency. It seems that neither low nor high phone density facilitated the correct identification of a token. Instead, mid-density was most favorable for perception.

The outcomes for token duration (expressed as phone density) provide some explanation for the effects of context, at least for reaction time; the other way to shed more light on non-native perception is to consider typological differences between L1 and L2 as it is very likely that typology plays a role. Out of six casual speech processes present in English (/t, d, h/ deletion, fricativization, place assimilation, and Yod coalescence), Polish has only two: /t/ deletion and assimilation of place (and of manner and voicing, Wierzchowska, 1980; Sawicka, 1995; Madelska, 2005). One may assume that the processes, which both languages have in common, will be the most salient for learners' perception (Table 4).

Below is an illustration of the correct perception of reduced forms across reduction categories in the form of Figure 4 (shown for the ContextGroup only since the NoContext group exhibited poor perception in general).

Regarding descriptive statistics, M = 403,5 SD = 95.46 for learners, and M = 574, SD = 143.68 for native speakers. Figure 4 reveals that for /h/ deletion, a process that is absent from Polish phonology (Table 4), perception was at the lowest level of 27%. The outcomes for fricativization (not present in Polish) and place assimilation (present in Polish) are the same, i.e., 38% of correct recognition, yielding mixed results in explaining perception with typological differences. /t/

TABLE 3 Phone density across groups.

NoContext	Number of neighboring phones	Example	Context	Number of neighboring phones	Examples
3–4	Low	*I had*	3–5	Low	*Did you get her*
7	Mid	*Last year*	15–16	Mid	*I still see her occasionally*
10–11	High	*childminder*	25–39	High	*I could not turn him on my own that is why I always had to ask for help*

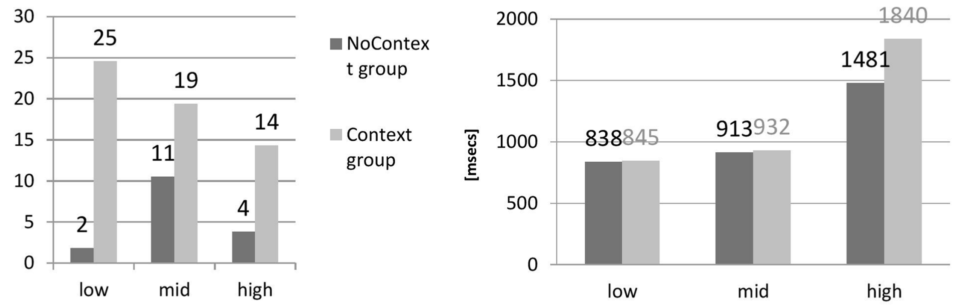


FIGURE 3
Correct answers across density groups (accuracy, unit: number of phonemes) and reaction time.

TABLE 4 Typological differences between Polish and English.

Process	Polish	English
Fricativization	–	+
Yod coalescence	–	+
Assimilation of place	+	+
h-deletion	–	+
d-deletion	–	+
t-deletion	+	+

deletion, another reduction process common in L1 and L2, was identified on nearly the same level of accuracy as /d/ deletion (not present in Polish) and less accurately than Yod coalescence (also not present in Polish). Should typological differences matter, we would expect place assimilation and /t/ deletion to rank the highest in terms of correct identification, followed by phonological processes not known in L1. Figure 4 clearly demonstrates that this, however, was not the case and undermines the influence of L1 on L2 in the area of reduced forms.

3.2 Results for hypothesis 2

Hypothesis 2 predicted that the subjects perceive words affected by processes of categorical reduction more accurately than the gradient ones. This prediction was made on the grounds that categorical reduction deletes a sound without an acoustic trace, and the difference between citation and reduced form is consequently more perceptually salient than a change of one sound's place on articulation into another (gradient reduction). Figure 5 displays the outcomes of the analysis.

Raw percentages from Figure 5 show that the perception of gradient and categorical reduction was barely different for the two groups across two conditions. Nevertheless, the difference estimated by one-way ANOVA between groups trended in various directions. For categorical reduction, they are reported as follows: NoContext Group learners, $F(1,101) = 70628.00$, $p < 0.05$, M = 480.67, SD = 190, N = 4,680, and Context Group learners: $F(1,4) = 0.006$, $p > 0.05$, M = 401.67, SD = 117.14, N = 3,000. However, the effect size was very weakly significant, $r = 0.24$, $d = 0.5$, for the NoContext vs. Context Condition. Concerning gradient reduction, this is how effect size looked for learners: the NoContext group: $F(1,101) = 334490.58$, $p < 0.05$, $r = 0.46$, $d = 1.04$ (the NoContext group: M = 442.33, SD = 147.29, N = 4,160; the Context Group: M = 259.40, SD = 201.17, N = 2,850). Thus, the effect size points to a medium (bordering weak) difference in gradient reduction for Polish learners of English.

The difference for native speakers was observed, $F(1,101) = 784188.32$, $p < 0.05$, for categorical reduction (NoContext: M = 624, SD = 222.14, N = 4,680, Context: M = 877.78, SD = 37.81, N = 3,000). The effect size, $r = -0.62$, $d = -1.59$, pointed to a medium-strong influence of the context on recognition. Regarding gradient reduction, the effect of a process type was robust, with $r = -0.78$, $d = -2.50$, $F(1,101) = 1567800.14$, $p < 0.05$ (NoContext: M = 537.33, SD = 39.72, N = 4,160, Context: M = 820.83, SD = 155.29, N = 2,850).

Following the effect size of the difference between categorical and gradient reduction (rather than mere values from Figure 5), the study reports that it was very weak in the no-context condition and weak in the context condition within the learners' group, yielding the effects statistically insignificant. Within the group of native users of English, on the other hand, the difference between categorical (medium strong) and gradient (strong) reduction shows that the prediction of Hypothesis 2 was correct. Native speakers perceived categorical reduction more accurately than the gradient one. Similar to Hypothesis 1, Hypothesis 2 is thus corroborated for native speakers and rejected for learners.

3.3 Results for hypothesis 3

The study hypothesized that learners with musical backgrounds perceive reduced forms better than those without. Figure 6 visualizes the obtained results.

Two-way ANOVA compared two groups of learners (learners with musical background, learners without musical background) in two conditions (with context and without context), revealing no difference in accuracy: for treatment (music), $F(1,101) = 3.96$, $p = 0.19$, and for other factors (context), $F(2,101) = 50.52$, $p = 0.02$; NoContext Group musical background, M = 37.48, SD = 7.13, N = 27, NoContext group no musical background, M = 35.67, SD = 7.45, N = 25, Context Group musical background: M = 48.70, SD = 14.14, N = 18, and Context group no musical background: M = 48.27, SD = 10.99, N = 32. The differences in terms of effect size were as follows: for context effects, $d = 0.02$, CI: from −0.52 to 0.56, $v = 0.077$. For music effects, $d = 0.00$, CI: from −0.57 to 0.58, $v = 0.09$. The results for the effects of musical background do not lend support to hypothesis 3 since the difference in accuracy

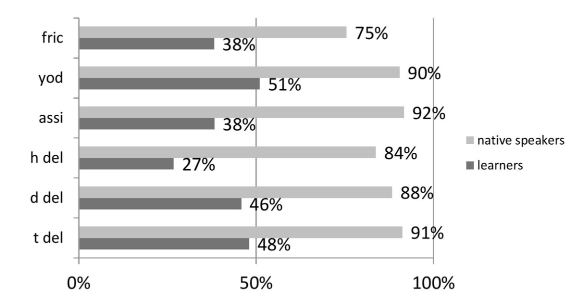


FIGURE 4
Distribution of accuracy across groups of reduced forms.

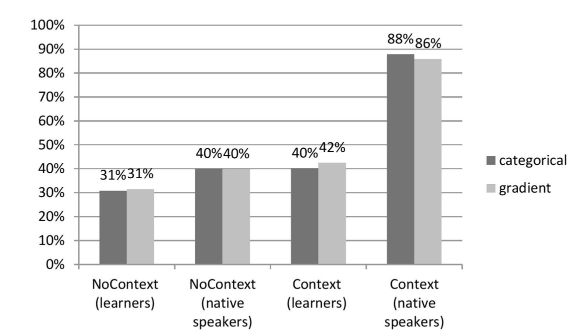


FIGURE 5
Distribution of accuracy within reduction categories across groups.

between the two groups (music) and condition (context) failed to reach statistical significance, with d not approximating the threshold of 1.

With respect to RT, the differences between groups were as follows: $F(2,101) = 128.040$, $p = 0.008$ (NoContext Group musical background: M = 1367.46, SD = 546.52, $N = 27$; NoContext group no musical background: M = 1581.55, SD = 660.16, $N = 25$; Context Group musical background: M = 1399.79, SD = 655.24, $N = 18$; Context group no musical background: M = 1336.54, SD = 860.24, $N = 32$). The differences in terms of effect size were as follows: for context effects, $d = -0.35$, CI: from −0.96 to 0.246, v = 0.09. For musical background effects, $d = 1.00$, CI: from 0.4528 to 1.56, v = 0.08. Thus, context effects were weak in terms of effect size but quite robust for musical background.

It has to be noted that, unlike for accuracy, there was a difference between the two groups with regard to reaction time ($d < 1.0$). Thus, it appears that the subjects with musical backgrounds recognized reduced forms faster but not more accurately in comparison with subjects who received no such training. In light of the outcomes, hypothesis 3 assumes a more accurate and faster perception of reduced forms by learners with musical backgrounds in comparison with learners with no such background, which is negatively verified for accuracy and positively for reaction time.

4 Discussion

The study was geared to furnish answers to three research questions related to the acquisition of reduced forms by non-native speakers of English. Specifically, hypothesis 1 tested the effects of lexical context and was positively verified for native users only. Similarly, hypothesis 2, exploring the effects of a process type, proved correct for native speakers of English and showed weak or even very weak effects for learners of English. The effects of musical background, hypothesis 3, were not found in accuracy and were present in reaction time. These outcomes are discussed in turn.

We know that, in general, learners display a reliance on semantic context (e.g., Van De Ven et al., 2011) and that they require much more contextual information than native users (Nooteboom and Truin, 1980; Koster, 1987). In a perception study, Pearman (2004) also found that "learners required more acoustic information […] than natives in order to identify a word often needing to hear the beginning of the next word before recognizing an item correctly" (Pearman, 2004, p. 1). The present study, however, cannot provide support for the role of context for learners from analysis of reduced forms. The finding regarding the effects of semantic context is that it fails to improve the perception of reduced forms in SLA. This suggests that unlike other aspects of language acquisition, processes of casual speech pose considerable processing difficulties for learners, and the study argues that the inclusion of semantic context alone does not suffice to perceive reduced forms correctly by learners.

Although Polish learners of English were not sensitive to the effects of lexical context in recognizing reduced forms, they were influenced by the token duration and phone density. Phone density was calculated as the number of lexical items in the phrase or sentence that contained a reduced form. As a result, three groups emerged: low-density (between 3 and 5 words in the Context condition), mid-density (ranging from 15 to 16 words surrounding the reduced form), and high-density (25–39 lexical items preceding and following the reduced form in question). Low density was expected to facilitate correct identification, whereas high density, associated with increased cognitive load for learners, should prove difficult for perception. Figure 3 clearly points to a link between the number of words (token duration) and reaction time in both Context and NoContext groups of learners. In addition, within the Context group, phone density had a robust effect on learners whose perception deteriorated as phone density grew. Some effects of phone density are also visible in the NoContext group of learners, where too-short or too-long tokens were not correctly identified. Interestingly, mid-phone density had the strongest effect on perception without context. Following the results of hypothesis 1, the study thus argues that for Polish learners of English, the effects of phone density are stronger than the effects of lexical context.

Apart from undermining the role of lexical context for learners assumed in previous studies (its effects were observed for native users), the outcomes of the study throw certain doubts on the importance of typological differences. Its results are not in line with the findings of Shockey and Bond (2012) and Shockey and Ćavar (2013). They found that Polish learners recognized reduction processes very well and that they were well equipped to identify reduced forms in English, thanks to the high frequency of reduction processes in Polish. Figure 4 however, does not lend support to typological differences' potential to explain learners' performance. The lack of conformity between the results of the present study and the abovementioned ones might stem from methodological issues: Shockey and Bond (2012) and Shockey and Ćavar (2013) used read speech. This study employs a speech corpus of casual English. The two studies focused on those processes which perform consonantal cluster reduction, whereas the present study was more varied in the selection

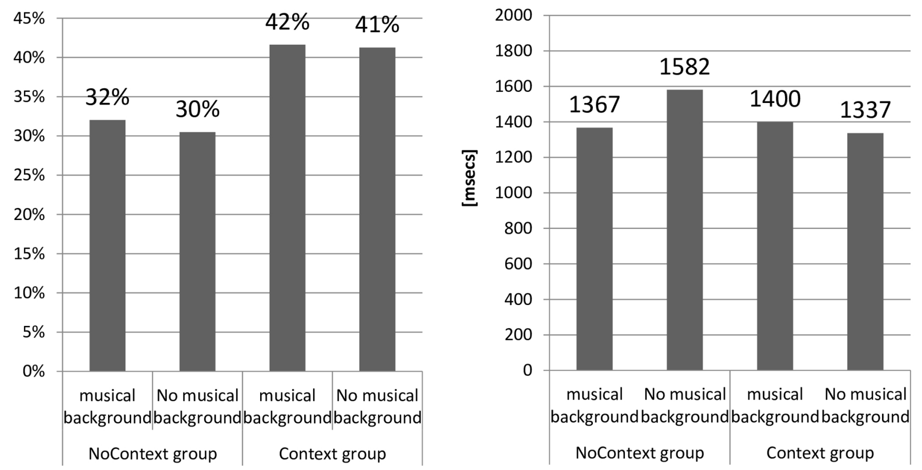


FIGURE 6
Acc and RT according to musical background.

of processes (e.g., fricativization, Yod coalescence, and assimilation of place are non-reductory processes). Overall, the outcomes of the two processes that English and Polish have in their phonologies were mixed. Should typological differences be really significant, assimilation of place and t-deletion would be recognized in a more accurate fashion than the remaining four, more typologically distant processes. In light of the results (Table 4), this was not necessarily the case. For individual sounds, the Perceptual Assimilation Model (Best and Tyler, 2007) predicts that a category of a sound from L2 is assimilated to the nearest or equivalent category in L1. Categories ascribed to, say, a single vowel sound are usually two-dimensional within one phonemic unit (frontness and height).

Processes of casual speech are much more complex than individual sounds. A straightforward assimilation of a process in the first and second languages does not do justice to the complexity of reduction processes which involve, among others, a class of consonants affected by the processes, the impact of preceding and following sounds (place assimilation, Yod coalescence, and fricativization), position in a word (e.g., /t, d/ deletion), morphology (Yod coalescence), amount of stress and the grammatical status of a word (/h/ deletion), and many other variables. Thus, the study proposes that the transfer of phonological process categories present in L1 to L2 might be governed by entirely different mechanisms than those governing vowels or consonants. It is suggested that the transfer mechanisms may not fully account for processes of causal speech due to their structural dependence on larger units and spanning word boundaries. A similar conclusion was reached by Bradlow and Pisoni (1999): "In other words, spoken language processing relies on both accurate phoneme categorization and knowledge of the sound structure of the target language" (Bradlow and Pisoni, 1999, p. 2,084).

Another argument speaks in favor of the perceptual difficulty of reduced forms, which goes beyond the assimilation of a category. Vowel reduction, unlike consonantal reduction, is a well-established language universal (Greenberg, 1978) and is governed by grammar, stress assignment, and vowel quality, among other factors. Consonantally reduced forms, the subject of the study, have much less uniform patterns than vowel reduction. Some consonants become elided /t, d, h/, while others assimilate to their contextual neighbors (Yod coalescence and place assimilation) or to lenis consonants (fricativization). In addition, reduced forms occur both within and across word boundaries; sometimes, they become fully lexicalized (e.g., Yod coalescence in *gradual*). The latter, according to the Production Planning Hypothesis (Kilbourn-Ceron et al., 2020), have a wide array of constraints, such as conditional probability or phonological context, which may well add to the difficulties in recognition. This suggests that poor perception of reduced forms by Polish learners can possibly be accounted for by their high structural complexity which surfaces as perceptual difficulties.

There is also the issue of the frequency of occurrence of a process that is connected to linguistic typology. Maddieson (1984) found that in his database of 451 languages of the world, the CV structure is preferred: over 70% of world languages exhibit such preference and have no consonant clusters. In this connection, processes such as consonant cluster reduction are much more frequent in Polish than in English, and this exposure to high-frequency processes might at least in part explain the results obtained in the course of verification of hypothesis 1. The study makes a suggestion that the frequency of a process in L1 and L2 could be a good predictor for perception and perhaps would provide a reliable explanation of non-native perception. This suggestion remains speculative since research on casual speech is qualitatively oriented at the expense of the quantitative aspect, i.e., statistics on the frequency of realization of processes. For instance, Shockey (1974) states that "this process [place assimilation] is quite frequent in connected speech" (Shockey, 1974, p. 36). Lack of precision in these statements has already been pointed out by Watson (2006): "statements such as 'plosives are frequently realized as affricates or fricatives' are common in the literature, but it is never clear exactly how frequent frequent is" (Watson, 2006, p. 150).

The present study finds no effect of lexical context for learners, nor can it fully explain their perception using language-specific differences, especially if they are treated in binary terms of presence or absence of a process causing a reduced form. Alternatively, lack of exposure to casual English and difficulty with unfamiliar accents are proposed here as possible sources of influence. Typically, pronunciation courses in an L2 classroom rely heavily on sound articulation at the expense of listening activities. Drilling vowels and consonants takes up so much of the syllabus that very little time, if any at all, is devoted to perception. The use of a speech corpus in a classroom is also an infrequent practice. On top of that, there is little exposure to authentic casual speech: one may speculate that the instructors adapt their speech to language learners and extensively employ citation forms.

There is also a recurring issue with the pronunciation standard for SLA. Lancashire, a non-standard variety of English used in the study, is a northern English dialect. It exhibits the following features: monophthongization, t-to-r realization, /l/ vocalization and Yod dropping which label the dialect "non-standard" compared to the standard variety (i.e. SSBE) (Orton and Halliday, 1962; Wells, 1982; Beal, 2008). Therefore, students of English might experience problems in recognizing individual words on the one hand (though frequency counts prove that the words were not difficult, Appendix 1). On the other hand, for native speakers who represented various English accents (Table 2), Lancashire (with context) did not hinder understanding. We know that even very young children can recognize the meaning of words in non-native accents: Mulak et al. (2013) demonstrated that 19-month-old Australian children could identify words in Jamaican English, linking this to specific language skills rather than overall cognitive ability. In addition, some subjects of the study from the learners' group indicated in the pre-study questionnaire that they had been working in the Northern UK and might have some familiarity with Lancashire or a neighboring dialect.

Turning to the effects of process type (hypothesis 2), the study reports very weak effects for the categorical type and weak effects for the gradient type within the learners group. On the other hand, the effects were significant for native speakers of English, which is in line with previous findings (Holst and Nolan, 1995; Hanique et al., 2013). This hypothesis was put forward because in articulation studies, processes that are complete (the categorical type) pattern in a different way than the incomplete, transient ones (the gradient type of phonological processes). Hypothesis 2 extends the famous stop–fricative metaphor to perception: "it is easier to run into a wall than to halt an inch in front of it" (source unknown, in Boersma, 1990) and assumes that categorical processes are more perception-friendly than the gradient ones. The study predicted that the loss of a segment (/t, d, h/ deletion) is perceptually more salient than the change of a sound into a new quality (place assimilation, Yod coalescence, and fricativization). According to Figure 5 , the perception of gradient and categorical reduction proved to be equally difficult for learners, whereas no differences between process types were reflected in the perception of native speakers. In light of these results, it seems that the type of process (gradient vs. categorical) plays no role in the perception of reduced forms in SLA as far as Polish is concerned. This might suggest that SL listeners do not perceive the fine-grained phonetic detail of a process.

Following previous studies, the present study extended the influence of musical background on the perception of reduced forms regarding consonants (hypothesis 3) and found no such effects for second language learners. Although the effects of musical background were attested in reaction time, they did not exert any influence on the accuracy of non-native perception. Contrary to these findings, studies by Milovanov (2009), Salcedo (2010), Besson et al. (2011), Gordon et al. (2011), and Balčytytė-Kurtinienė (2018) found the effects of musical training and aptitude on the perception of rhythm, pitch, and vowel reduction. It has to be noted that the reduction of vowels has the potential to change the rhythm and the syllabic structure of vowels, to which people with musical talent have reacted as expected. The reduction caused by consonants, on the other hand, is not capable of changing rhythmic properties or general melody of an utterance (with a notable exception of h-deletion in an unstressed pronominal form, which is still strongly affected by associated vowel change and stress shift). The lack of effects of musical background on the non-native perception of reduced forms, reported in the study, can be linked to the generally low melodic impact of consonantal processes on speech.

Another noteworthy observation is that RTs were longer for native speakers in comparison with L2 learners in the context condition (Figure 2). The speed–accuracy trade-off (Standage et al., 2014; Huang et al., 2017) explains that responses are more accurate and slower if the experimental conditions focus on accuracy. If the conditions favor speed, the responses are less accurate and faster. The study's instructions had no conditions highlighting either accuracy or speed. There was no time limit for typing the responses or pressing a Yes/No button. In addition, the instruction did not stress that participants should type exactly what they hear, nor should they type the full phrase. The lack of clear experimental focus could have possibly given rise to more variability in the obtained results, both for learners and the control group of native speakers.

The study addresses non-native perception of reduced forms using a wide variety of consonantal processes from a corpus of casual speech and fills the gaps identified in the Introduction section. The study tests the effects of lexical context, type of process, and musical background, drawing a number of conclusions and making a number of novel contributions to the field of SL perception. The first conclusion is that hypothesis 1 positively verified the effects of lexical context for the native perception of reduced forms but not for the learners. The study suggests that learners used cues from phone density instead of lexical context or typological differences between L1 and L2. There is also the possibility that language-specific differences might surface more vividly if the frequency of process occurrence is considered. Drawing on exposure to reduced forms, current classroom practice, and perhaps lack of familiarity with various dialects of English can be enumerated as additional factors shaping perception. The study speculates that the general complexity of phonological processes affecting consonants calls for exploring the role of awareness and metacompetence. The second conclusion is that the effects of a process type are insignificant for learners and robust for native users. It points to a greater role of phonetic details in the recognition of phonological processes than expected based on production studies. The final conclusion is that there are no effects of musical background on the perception of consonantal reduction for learners.

In general, the study contributes to linguistic theory by increasing our understanding of the effects governing non-native perception of casual speech. Of particular theoretical interest in this study is the overall implication that, apparently, the perception of reduced forms in a second language might proceed according to entirely different mechanisms than those governing the perception of vowels or consonants. In general, the effects tested (i.e., context, process type, and musical background) suggest that the role of awareness in the context of casual speech is greater than that of segments. Perhaps, reduced forms exist as independent representations in non-native phonology, not accessed as a deviation from citation form but as separate entities. The results of the study raise new research questions and prove that the perception of reduced forms is well worth pursuing further and deserves a more exhaustive treatment. It has to be noted that recognition of reduced forms is much more perceptually challenging than recognition of individual sounds or words as has already been demonstrated by the results obtained in this study for native and non-native speakers of English.

From a methodological viewpoint, the study demonstrates the utility of a speech corpus (PAC, Durand and Pukli, 2004) for investigating the recognition of reduced forms. The study, however, is not devoid of limitations. First, it tested the perception of only one variety of English, which happened to be a non-standard one. Second, the corpus used in the study had only nine speakers and totaled about 4.5 h of casual speech, which is a relatively small sample to mine reduction processes. The sample size could be enlarged and more varied in terms of speakers. The third shortcoming of the study might possibly stem from the study design itself: no data were accessible from the E-prime script, whether reaction time and accuracy come from the first or second hearing.

These limitations serve to outline several implications for further studies. One of them is to conduct a cross-linguistic investigation of reduced forms in a number of typologically related and unrelated languages (preferably with parallel reduction processes) to determine the role of typological differences or similarities. Another implication is to tease apart the effects of process type under carefully constructed experimental conditions. A possible research venue is to include accents other than Lancashire with a view of substantiating any claims about Lancashire's difficulty or unfamiliarity and answering the question of to what extent processes of connected speech are dialect-specific. In addition, future studies might pursue the same research questions using an entirely different pool of subjects, such as learners of English for communicative purposes, immigrants, and professional musicians. Finally, the use of an eye-tracking research paradigm could take non-native perception to a new level.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary material, further inquiries can be directed to the corresponding author.

Ethics statement

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

MK: Writing – original draft, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. The author gratefully acknowledges the financial support of the National Science Center (grant number 2012/05/D/HS2/03565).

Conflict of interest

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2024.1305134/full#supplementary-material

References

Balčytytė-Kurtinienė, G. (2018). The music of English connected speech: how and why singing affects EFL learners pronunciation. Paper presented at the accents 2018 conference, Łódź, Poland. Available at: http://filolog.uni.lodz.pl/accents/Accents%20MMXVIII%20-%20abstracts.pdf (Accessed 27 February 2019).

Barry, M. (1992). Palatalisation, assimilation, and gestural weakening. *Speech Comm.* 11, 393–400. doi: 10.1016/0167-6393(92)90044-8

Beal, J. (2008). "English dialects in the north of England: phonology" in *Varieties of English, volume 1*. eds. B. Kortman and C. Upton (Berlin: Mouton de Gruyter), 122–144.

Beal, J. (2010). *An introduction to regional Englishes*. Edinburgh: Edinburgh University Press.

Bell, A. (1984). Language style as audience design. *Lang. Soc.* 13, 145–204. doi: 10.1017/S004740450001037X

Bell, A. (2001). "Back in style: reworking audience design" in *Style and sociolinguistic variation*. eds. P. Eckert and J. R. Rickford (Cambridge: CUP), 139–169.

Besson, M., Chobert, J., and Marie, C. (2011). Transfer of training between music and speech: common processing, attention, and memory. *Front. Psychol.* 2, 1–12. doi: 10.3389/fpsyg.2011.00094

Best, C. T., and Tyler, M. D. (2007). "Nonnative and second-language speech perception: commonalities and complementarities" in *Language experience in second language speech learning: In honor of James Emil Flege*. eds. M. J. Munro and O.-S. Bohn (Amsterdam: John Benjamins), 13–34.

Boersma, P. (1990). Modelling the distribution of consonant inventories. In: *Paper presented at the Congress of Linguistics and Phonetics*, 30 August 1990, Prague, the Czech Republic). Available at: http://www.hum.uva.nl/paul/papers/Praag_1990.pdf (Accessed 3 January 2021).

Bosch, L.ten, Giezenaar, G., Boves, L., and Ernestus, M. (2016). Modeling language-learners' errors in understanding casual speech. In G. Adda, V. Barbu Mititelu, J. Mariani, D. Tufiş and I. Vasilescu (Eds.), *Errors by humans and machines in multimedia, multimodal, multilingual data processing*. Bucharest: Editura Academiei Române. 07–121.

Bradlow, A. R., and Pisoni, D. B. (1999). Recognition of spoken words by native and non-native listeners: talker-, listener- and item-related factors. *J. Acoust. Soc. Am.* 106, 2074–2085. doi: 10.1121/1.427952

Brand, S., and Ernestus, M. (2018). Listeners' processing of a given reduced word pronunciation variant directly reflects their exposure to this variant: evidence from native listeners and learners of French. *Q. J. Exp. Psychol.* 71, 1240–1259. doi: 10.1080/17470218.2017.1313282

British National Corpus. (2017). Available at: http://corpus.byu.edu/bnc/

Browman, C. P., and Goldstein, L. (1990). "Tiers in articulatory phonology, with some implications for casual speech" in *Papers in laboratory phonology I: Between the grammar and physics of speech*. eds. J. Kingston and M. E. Beckman (Cambridge: Cambridge University Press), 341–376.

Bussmann, H. (1996). *Routledge dictionary of language and linguistics*. London and New York: Routledge.

Bybee, J. L. (2013). "Usage-based theory and exemplar representation" in *The Oxford handbook of construction grammar*. eds. T. Hoffman and G. Trousdale (Oxford: Oxford University Press), 49–69.

Clements, G. N. (1985). The geometry of phonological features. *Phonol. Yearb.* 2, 225–252. doi: 10.1017/S0952675700000440

Cruttenden, A. (2008). *Gimson's pronunciation of English*. *7th* Edn. London: Hodder.

Dilley, L. C., and Pitt, M. A. (2010). Altering context speech rate can cause words to appear or disappear. *Psychol. Sci.* 21, 1664–1670. doi: 10.1177/0956797610384743

Dunaj, B. (1985). *Grupy spółgłoskowe współczesnej polszczyzny mówionej [Consonant clusters in contemporary spoken Polish]*. Warszawa: Polska Akademia Nauk.

Dunaj, B. (2006). Zasady poprawnej wymowy polskiej [Rules of Polish pronunciation]. *Język polski* LXXXVI, 161–172.

Durand, J., and Pukli, M. (2004). How to construct a phonological corpus: PRAAT and the PAC project. *Trib Int Langues Vivantes* 36, 36–46.

Ellis, L., and Hardcastle, W. (2002). Categorical and gradient properties of assimilation in alveolar to velar sequences: evidence from EPG and EMA. *J. Phon.* 30, 373–396. doi: 10.1006/jpho.2001.0162

Ernestus, M., Kouwenhoven, H., and van Mulken, M. (2017). The direct and indirect effects of the phonotactic constraints in the listener's native language on the comprehension of reduced and unreduced word pronunciation variants in a foreign language. *J. Phon.* 62, 50–64. doi: 10.1016/j.wocn.2017.02.003

Ernestus, M., and Warner, N. (2011). An introduction to reduced pronunciation variants. *Editor J Phonet* 39, 253–260. doi: 10.1016/S0095-4470(11)00055-6

Fasold, R. (1972). *Tense marking in black English*. Arlington, VA: Center for Applied Linguistics.

Gordon, R. L., Magne, C. L., and Large, E. W. (2011). EEG correlates of song prosody: a new look at the relationship between linguistic and musical rhythm. *Front. Psychol.* 2:352. doi: 10.3389/fpsyg.2011.00352

Gow, D. W. (2003). Feature parsing: feature cue mapping in spoken word recognition. *Percept. Psychophys.* 65, 575–590. doi: 10.3758/BF03194584

Greenberg, J. H. (1978). *Universals of human language, Volume 2: Phonology*. Stanford: Stanford University Press.

Gut, U., and Voormann, H. (2014). "Corpus design" in *The Oxford handbook of corpus phonology*. eds. J. Durand, U. Gut and G. Kristoffersen (Oxford: OUP), 13–26.

Guy, G. R. (1980). Variation in the group and the individual: The case of final stop deletion. In Locating language in time and space. Academic Press. 1–36.

Guy, G. R. (1992). Explanation in variable phonology: an exponential model of morphological constraints. *Lang. Var. Chang.* 3, 1–22. doi: 10.1017/S0954394500000429

Guy, G. R., and Boberg, C. (1997). Inherent variability and the obligatory contour principle. *Lang. Var. Chang.* 9, 149–164. doi: 10.1017/S095439450000185X

Hanique, I., Ernestus, M., and Schuppler, B. (2013). Informal speech processes can be categorical in nature, even if they affect many different words. *J. Acoust. Soc. Am.* 133, 1644–1655. doi: 10.1121/1.4790352

Harvey, A. (2017). *Music, evolution and the harmony of souls*. Oxford: Oxford University Press.

Hazen, K. (2011). Flying high above the social radar: coronal stop deletion in modern Appalachian. *Lang. Var. Chang.* 23, 105–137. doi: 10.1017/S0954394510000220

Holst, T., and Nolan, F. (1995). "The influence of syntactic structure on [s] to [♠] assimilation" in *Phonology and phonetic evidence: Papers in laboratory phonology IV*. eds. B. Connell and A. Arvaniti (Cambridge: Cambridge University Press), 315–333.

Huang, J., Rathod, V., Sun, C., Zhu, M., Korattikara, A., Fathi, A., et al. (2017). Speed/accuracy trade-offs for modern convolutional object detectors. In: *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 7310–7311).

Jackendoff, R. (2009). Parallels and nonparallels between language and music. *Music Percept Interdisciplin J* 26, 195–204. doi: 10.1525/mp.2009.26.3.195

Johnson, K. (2004). "Massive reduction in conversational American English". In: Yoneyama, K., and Maekawa, K.. (eds.) *Spontaneous speech: Data and analysis. Proceedings of the 1st session of the 10th international symposium*. Tokyo, Japan: The National International Institute for Japanese Language. 29–54.

Kilbourn-Ceron, O., Clayards, M., and Wagner, M. (2020). Predictability modulates pronunciation variants through speech planning effects: a case study on coronal stop realizations. *Lab Phonol* 11:5. doi: 10.5334/labphon.168

Kortmann, B., and Upton, C. (2008). *Varieties of English, Volume 1*: Berlin: Mouton de Gruyter.

Koster, C. (1987). *Word recognition in foreign and native language*. Dordrecht, The Netherlands: Foris.

Labov, W. (1994). *Principles of language change: Internal factors*. Oxford: Blackwell.

Labov, W. C., Robins, C., and Lewis, J. (1968). A study of the non-standard English of the negro and Puerto Rican speakers in new York City (vol. 1). *Final research report, Cooperative research project no 3288*. Washington, DC: US Office of Education.

Lange, M. V., and Laganaro, M. (2014). Inter-subject variability modulates phonological advance planning in the production of adjective-noun phrases. *Front. Psychol.* 5:43. doi: 10.3389/fpsyg.2014.00043

Lindblom, B. (1963). Spectrographic study of vowel reduction. *J. Acoust. Soc. Am.* 35, 1773–1781. doi: 10.1121/1.1918816

Lodge, K. (1984). *Studies in the phonology of colloquial English*, Kent: Croon Helm.

Maddieson, I. (1984). *Patterns of sounds*. Cambridge: CUP.

Madejowa, M. (1987). Współczesna polska norma językowa w zakresie wymowy samogłosek nosowych [Contemporary Polish linguistic norm of nasal vowels]. *Studia Phonetica Posnaniensa* 1, 31–56.

Madejowa, M. (1993). Normative rules of modern polish pronunciation. *Studia Phonetica Posnaniensa* 4:294.

Madelska, L. (2005). *Słownik wariantywności fonetycznej współczesnej polszczyzny [Dictionary of phonetic variability in contemporary Polish]*. Kraków: Collegium Columbinum.

Magne, C., Schon, D., and Besson, M. (2006). Musician children detect pitch violations in both music and language better than nonmusician children: Behavioural and Electrophysio-logical approaches. *J. Cogn. Neurosci.* 18, 199–211. doi: 10.1162/jocn.2006.18.2.199

McCarthy, J. (1988). Feature geometry and dependency: a review. *Phonetica* 45, 84–108. doi: 10.1159/000261820

Milovanov, R. (2009). The connectivity of musical aptitude and foreign language learning skills: neural and Behavioural evidence. *Anglicana Turkuensia* 27, 1–56.

Milovanov, R., Pietilä, P., Tervaniemi, M., and Esquef, P. A. A. (2010). Foreign language pronunciation skills and musical aptitude: a study of Finnish adults with higher education. *Learn. Individ. Differ.* 20, 56–60. doi: 10.1016/j.lindif.2009.11.003

Mithen, S. (2005). *The singing Neanderthals: The origins of music, language, mind and body*. London: Weidenfeld & Nicholson.

Mitterer, H., and Ernestus, M. (2006). Listeners recover/t/s that speakers reduce: evidence from/t/−lenition in Dutch. *J. Phon.* 34, 73–103. doi: 10.1016/j.wocn.2005.03.003

Mitterer, H., Yoneyama, K., and Ernestus, M. (2008). How we hear what is hardly there: mechanisms underlying compensation for/t/−reduction in speech comprehension. *J. Mem. Lang.* 59, 133–152. doi: 10.1016/j.jml.2008.02.004

Mora, C. F., and Gant, M. (2016). *Melodies, rhythm and cognition in foreign language learning*. Cambridge: Cambridge Scholars Publishing.

Morano, L., Ten Bosch, L., and Ernestus, M. (2019). "Looking for exemplar effects: testing the comprehension and memory representations of r'duced words in Dutch learners of French" in *Speech perception and production: Learning and memory*. eds. S. Fuchs, A. Rochet-Capella and J. Cleland (Bern: Peter Lang Publisher)

Mulak, K., Best, C., and Tyler, M. (2013). Development of phonological constancy: 19-month-olds, but not 15-month-olds, identify words in a non-native regional accent. *Child Dev.* 84, 2064–2078. doi: 10.1111/cdev.12087

Neu, H. (1980). "Ranking of constraints on /t, d/ deletion in American English: a statistical analysis" in *Locating language in time and space*. ed. W. Labov (New York: Academic Press), 37–54.

Nolan, F. (1992). "The descriptive role of segments: evidence for assimilation" in *Papers in laboratory phonology II: Gesture, segment, prosody*. eds. D. R. Ladd and G. J. Docherty (Cambridge: Cambridge University Press), 261–280.

Nooteboom, S.& Truin, P. (1980). Word recognition from fragments of spoken words by native and non-native listeners. IPO Annual Progress Report, 15, 42–47.

Nouveau, D. (2012). Limites perceptives de l'e caduc chez des apprenants néerlandophones. *Revue Canadienne de linguistiquea Appliquée* 15, 60–78.

Ohala, J. J. (1990). "The phonetics and phonology of aspects of assimilation" in *Papers in laboratory phonology I: Between the grammar and the physics of speech*. eds. J. Kingston and M. Beckman (Cambridge: Cambridge University Press), 258–275.

Orton, H., and Halliday, W. J. (1962): *Survey of English dialects, Vol.1. (The six Northern countries and the Isle of Man)*. Leeds: E.J. Arnold & Son Ltd.

Orzechowska, P. (2019). *Complexity in polish phonotactics: On features, weights, rankings and preferences*. Singapore, Springer, 359.

Pastuszek-Lipińska, B. (2009). Do you want to pronounce correctly in a foreign language? Start musical lessons! *J. Acoust. Soc. Am.* 125:2755. doi: 10.1121/1.4784624

Patel, A. D. (2003). *Rhythm in language and music. Parallels and differences*. New York: New York Academy of Sciences.

Patel, A. D., and Daniele, J. R. (2003). An empirical comparison of rhythm in language and music. *Cognition* 87, B35–B45. doi: 10.1016/S0010-0277(02)00187-7

Pearman, A. (2004). Sound change in the learner: the perception of connected speech. In: *Paper presented at 9th Conference on Laboratory Phonology*, June 24–26, 2004.

Salcedo, C. S. (2010). The effects of songs in foreign language classroom on text recall, delayed text recall and involuntary mental rehearsal. *J Coll Teach Learn* 7, 19–30. doi: 10.19030/tlc.v7i6.126

Sawala, K., Szczegóła, T., and Weckwerth, J. (2024). *Say it right 5.0. A professional platform for learning English*. Available at: https://www.sayitrightonline.pl/pl-PL/Home

Sawicka, I. (1995). "Fonologia [phonology]" in *Gramatyka współczesnego języka polskiego. Vol. 3. Fonetyka i fonologia. [the grammar of the contemporary polish language. Vol. 3. Phonetics and phonology.]*. ed. H. Wróbel (Kraków: Wydawnictwo Instytutu Języka Polskiego PAN), 105–191.

Shockey, L. (1974). Phonetic and phonological properties of connected speech. Ohio State Working Papers in Linguistics (17). 1–143.

Shockey, L. (2003). *Sound patterns of spoken English*. Oxford: Blackwell.

Shockey, L., and Bond, Z. S. (2012). Holistic perception of phonological variants in rhythm, melody and harmony in speech. *Speech Lang Technol* 14, 199–209.

Shockey, L., and Ćavar, M. (2013). Roadrunners and eagles. *Res Lang* 11, 97–102. doi: 10.2478/v10015-012-0012-x

Smiljanic, A., and Bradlow, S. (2011). Bidirectional clear speech perception benefit for native and high-proficiency non-native talkers and listeners: intelligibility and accentedness. *J. Acoust. Soc. Am.* 130, 4020–4031. doi: 10.1121/1.3652882

Spinelli, E., and Gros-Balthazard, F. (2007). Phonotactic constraints help to overcome effects of schwa deletion in French. *Cognition* 104, 397–406. doi: 10.1016/j.cognition.2006.07.002

Standage, D., Blohm, G., and Dorris, M. C. (2014). On the neural implementation of the speed-accuracy trade-off. *Front. Neurosci.* 8:236. doi: 10.3389/fnins.2014.00236

Sumner, M. (2011). The role of variation in the perception of accented speech. *Cognition* 119, 131–136. doi: 10.1016/j.cognition.2010.10.018

Tagliamonte, S. A., and Temple, R. (2005). New perspectives on an ol' variable: (t,d) in British English. *Lang. Var. Chang.* 17, 281–302. doi: 10.1017/S0954394505050118

Tamminga, M. (2018). Modulation of the following segment effect on English coronal stop deletion by syntactic boundaries. *Glossa* 3, 1–27. doi: 10.5334/gjgl.489

Tamminga, M., MacKenzie, L., and Embick, D. (2016). The dynamics of variation in individuals. *Linguist Variation* 16, 300–336. doi: 10.1075/lv.16.2.06tam

Tanner, J., Sonderegger, M., and Wagner, M. (2017). Production planning and coronal stop deletion in spontaneous speech. *Lab Phonol* 8, 1–39. doi: 10.5334/labphon.96

University of Adam Mickiewicz. (n.d.). Available at: https://usosweb.amu.edu.pl/kontroler.php?_action=katalog2/przedmioty/pokazPrzedmiot&kod=15-PNJA-CP-1BA-12 (Accessed 27 January 2019).

Ven, M.van de, Ernestus, M., and Schreuder, R. (2012). Predicting acoustically reduced words in spontaneous speech: the role of semantic/syntactic and acoustic cues in context. *Lab Phonol*, 3, 455–481. doi: 10.1515/lp-2012-0020

Ven, M.Van De, Tucker, B. V., and Ernestus, E. (2011). Semantic context effects in the comprehension of reduced pronunciation variants. *Mem. Cognit.*, 39, 1301–1316, doi: 10.3758/s13421-011-0103-2

Van Dommelen, W., and Hazan, V. (2012). Impact of talker variability on word recognition in non-native listeners. *J. Acoust. Soc. Am.* 132, 1690–1699. doi: 10.1121/1.4739447

Viebahn, M., Ernestus, M., and McQueen, J. (2015). Syntactic predictability in the recognition of carefully and casually produced speech. *J. Exp. Psychol. Learn. Mem. Cogn.* 41, 1684–1702. doi: 10.1037/a0039326

Wagner, M. (2012). Locality in phonology and production planning. McGill Working Papers in Linguistics, 22, 1–18.

Warner, N. L., Brenner, D., Tucker, B. V., Sung, J. H., Ernestus, M., Simonet, M., et al. (2012). Processing reduced speech across languages and dialects. *J. Acoust. Soc. Am.* 132:1935. doi: 10.1121/1.4755115

Watson, K. (2006). The phonetics and phonology of plosive lenition in Liverpool English. PhD dissertation, Lancaster University.

Wells, J. (1982): *Accents of English*, Cambridge: Cambridge University Press.

Wierzchowska, B. (1980). *Fonetyka i fonologia języka polskiego [the phonetics and phonology of the polish language]*. Wrocław, Warszawa, Kraków, Poznań: Ossolineum.

Wolfram, W. (1969). *A sociolinguistic description of Detroit negro speech*. Washington, DC: Center for Applied Linguistics.

Wong, S. W., Mok, P. P., Chung, K. K. H., Leung, V. W., Bishop, D. V., and Chow, B. W. Y. (2017). Perception of native English reduced forms in Chinese learners: Its role in listening comprehension and its phonological correlates. *TESOL Quarterly* 51, 7–31.

Wright, S., and Kerswill, P. (1989). Electropalatography in the study of connected speech processes. *Clin. Linguist. Phon.* 3, 49–57. doi: 10.3109/02699208908985270

Zimmerer, F., and Reetz, H. (2014). Do listeners recover "deleted" final/t/ in German? *Front. Psychol.* 5, 73–103. doi: 10.3389/fpsyg.2014.00735

Zsiga, E. (1995). "An acoustic and electropalatographic study of lexical and postlexical palatalization in American English" in *Phonology and phonetic evidence: Papers in laboratory phonology IV*. eds. B. Connell and A. Arvaniti (Cambridge: Cambridge University Press), 282–302.

Zydorowicz, P. A. (2019). *Polish (mor)phonotactics in first language acquisition, connected speech and cluster processing*. Poznań, Wydawnictwo Naukowe Uniwersytetu im. Adama Mickiewicza, 361.

TYPE Original Research
PUBLISHED 24 July 2024
DOI 10.3389/fpsyg.2024.1295379

OPEN ACCESS

EDITED BY
Monika M. Polczynska,
University of California, Los Angeles,
United States

REVIEWED BY
Zuzanna Fuchs,
University of Southern California,
United States
Margarethe McDonald,
University of Kansas, United States

*CORRESPONDENCE
Alisa Baron
✉ barona@uri.edu

RECEIVED 16 September 2023
ACCEPTED 21 June 2024
PUBLISHED 24 July 2024

CITATION
Baron A, Connell K, Kleinman D, Bedore LM and Griffin ZM (2024) Grammatical gender in spoken word recognition in school-age Spanish monolingual and Spanish–English bilingual children. *Front. Psychol.* 15:1295379. doi: 10.3389/fpsyg.2024.1295379

Grammatical gender in spoken word recognition in school-age Spanish monolingual and Spanish–English bilingual children

Alisa Baron[1]*, Katrina Connell[2], Daniel Kleinman[3], Lisa M. Bedore[4] and Zenzi M. Griffin[5]

[1]Department of Communicative Disorders, University of Rhode Island, Kingston, RI, United States, [2]Independent Researcher, Florence, SC, United States, [3]Yale Child Study Center, Yale University, New Haven, CT, United States, [4]Department of Communication Sciences and Disorders, Temple University, Philadelphia, PA, United States, [5]Department of Psychology, University of Texas at Austin, Austin, TX, United States

This study examined grammatical gender processing in school-aged children with varying levels of cumulative English exposure. Children participated in a visual world paradigm with a four-picture display where they heard a gendered article followed by a target noun and were in the context where all images were the same gender (same gender), where all of the distractor images were the opposite gender than the target noun (different gender), and where all of the distractor images were the opposite gender, but there was a mismatch in the gendered article and target noun pair. We investigated 51 children (aged 5;0–10;0) who were exposed to Spanish since infancy but varied in their amount of cumulative English exposure. In addition to the visual word paradigm, all children completed an article–noun naming task, a grammaticality judgment task, and standardized vocabulary tests. Parents reported on their child's cumulative English language exposure and current English language use. To investigate the time course of lexical facilitation effects, looks to the target were analyzed with a cluster-based permutation test. The results revealed that all children used gender in a facilitatory way (during the noun region), and comprehension was significantly inhibited when the article–noun pairing was ungrammatical rather than grammatical. Compared to children with less cumulative English exposure, children with more cumulative English exposure looked at the target noun significantly less often overall, and compared to younger children, older children looked at the target noun significantly more often overall. Additionally, children with lower cumulative English exposure looked at target nouns more in the different-gender condition than the same-gender condition for masculine items more than feminine items.

KEYWORDS

grammatical gender, bilingualism, eye tracking, visual word paradigm, typically-developing children

1 Introduction

Children and adults process spoken language incrementally, utilizing the partial information at any given moment to predict upcoming words based on lexical or morphosyntactic cues, among others (e.g., Marslen-Wilson, 1987; Bates et al., 1996; Friederici and Jacobsen, 1999; Fernald et al., 2001). Grammatical gender is one example of such a cue that both children and adults efficiently use in spoken word recognition (or visual world paradigm) (Dahan et al., 2000; Lew-Williams and Fernald, 2007). The present study investigates spoken word recognition in school-aged children who speak Spanish and the influence of cumulative English exposure on this ability.

1.1 Spanish grammatical gender

Languages such as Spanish, French, and Dutch (and many others) assign grammatical gender to nouns. The gender of a given noun impacts the surface form of other words in the sentence that modify it, such as adjectives and determiners, so that they agree in gender and number. Spanish has two genders: masculine and feminine. Spanish nouns have a reliable pattern of overt gender marking, with masculine nouns ending in *–o* 99.9% of the time and feminine nouns ending in *–a* 96.6% of the time (Teschner and Russell, 1984). Although there are nouns that do not follow this reliable pattern, here, our focus is on nouns with canonical -o/-a endings. Definite and indefinite articles in Spanish are among the most frequently used words, and they agree in gender and number with the noun they modify. Singular masculine nouns are preceded by "*el*" or "*un*," and singular feminine nouns are preceded by "*la*" or "*una*." This regularity of the Spanish gender system facilitates early monolingual child acquisition of grammatical gender in Spanish articles (Perez-Pereira, 1991).

1.2 Acquisition and use of grammatical gender in monolinguals

For monolingually Spanish-exposed children, grammatical gender emerges in production at approximately 1;6 (year;months) (Hernández-Pina, 1984; Lleó, 1998), and they master gender agreement in production by age 3 (Soler, 1984; Perez-Pereira, 1991; López-Ornat, 1997; Eisenchlas, 2003; Castilla and Pérez-Leroux, 2010; Mariscal and Auza, 2017). Spanish–English bilinguals, Spanish speakers with a developmental language disorder (DLD), and Spanish heritage speakers produce more gender errors than their age-matched monolingual peers, which suggests that their acquisition differs from that of typically developing, monolingually-exposed children (Bruhn De Garavito and White, 2002; Bedore and Leonard, 2005; Morgan et al., 2013; Cuza and Pérez-Tattam, 2015).

In online language comprehension tasks, toddlers (2–3 years old) learning Spanish as their first language use morphosyntactic markers of grammatical gender to rapidly identify visual referents. Monolingual Spanish-speaking children take advantage of gender-marked words in real time to interpret spoken sentences rapidly (Lew-Williams and Fernald, 2007). In the study by Lew-Williams and Fernald (2007), children were presented with two objects while they listened to a speaker name one of the objects using an article + noun combination. Their task was to look at the picture that is named during a pre-recorded sentence (i.e., "Where's the [ball]? and Do you see it?"). The experiment included two conditions: a different-gender condition and a same-gender condition. For the same-gender trials, the two presented objects had the same grammatical gender, thus requiring the listener to wait for the noun to ascertain which object was being referred to. For the different-gender trials, the two presented objects differed in their grammatical gender, thus allowing the article to be predictive of the upcoming referent noun. Children's eye movements were tracked as they performed the task, and the results showed that these monolingual Spanish-speaking children oriented faster to the correct referent on different-gender trials where the article was informative, compared to the same-gender trials where the article was uninformative. This study indicated that adjacent, informative, grammatical cues (such as a gendered article) are used to facilitate online speech comprehension.

Similarly, preschool (5;4–6;6) and first grade (6;5–7;7) children in Chile were shown 32 distinct visual displays with four pictures of familiar objects at a time, each with a pre-recorded article–noun phrase matching one of the pictures (i.e., !Mira La manzana [Look! $\text{The}_{\text{ART.F.SG}}$ apple]) (Coloma et al., 2023). Researchers investigated whether children would use grammatical gender predictively in regard to definiteness (definite/indefinite), gender (masculine/feminine), and number (singular/plural). Children heard different combinations of definite (i.e., *el, la, los,* and *las*) and indefinite (i.e., *un, una, unos,* and *unas*) articles. These children tended to look more at the target object compared to the surrounding competitors even before the target object was mentioned (Coloma et al., 2023). In this longitudinal study, preschoolers showed a small but reliable anticipation effect to the target noun (~150 ms after the onset of the target noun) and, once in first grade, attended to the target noun ~300 ms before it was even mentioned. Within a similar design, bilingual Spanish–Catalan children (4;6–12;2) in Spain also saw four objects with conditions varying by definiteness, number, and gender with a pre-recorded sentence matching one of the objects (i.e., La chica muerde la manzana [The girl bites the apple]). Children were also able to identify the correct target before the target noun was stated based on the preceding gender-marked article (Christou et al., 2020).

1.3 Use of grammatical gender in bilingual children

Early bilingual acquisition of grammatical gender has been shown to be modulated by exposure and language dominance in simultaneous (exposure to both languages from birth) and sequential (exposure to a second language after age 3) bilingual children (e.g., Cornips and Hulk, 2008; Unsworth, 2013; Unsworth et al., 2014; Rodina and Westergaard, 2017). Studies on gender processing in bilingual children are quite sparse. Children (aged 8–9) who learn languages simultaneously, where both languages utilize grammatical gender, appear sensitive to grammatical gender cues (Lemmerth and Hopp, 2019). However, it appears that there are cross-linguistic influences at play and a possible influence of age of acquisition (i.e., simultaneous vs. sequential bilinguals). In another study of sequential Mandarin–Italian bilinguals, researchers found that a subset of the Mandarin–Italian bilinguals were slower or did not use grammatical gender predictively at all when compared to Italian-speaking monolingual peers (Bosch et al., 2022). They noted that gender processing was significantly affected by proficiency in the second language, Italian. Finally, in a study of 8- to 12-year-old early bilingual and multilingual children learning Italian in a majority context, there was a clear anticipatory gender effect, but multilingual children were slower than their Italian-speaking monolingual peers (Bosch and Foppolo, 2023).

In the US, many bilingual children of immigrant parents, who are referred to as heritage speakers, learn an ethnolinguistically minority language (Silva-Corvalán, 1994; Kondo-Brown, 2006). Within the US, many children who are heritage speakers of Spanish begin their formal education as Spanish-dominant speakers. However, these children quickly transition to become English-dominant due to their increased exposure to English in school. As time goes on and they progress in school (typically English only), they learn to read only in English, and the amount of input in Spanish decreases and becomes mostly limited to use at home (Montrul et al., 2008; Montrul, 2016). This decreased input and lack of literacy in Spanish can negatively impact these children's grammatical skills, leading to more grammatical errors (Montrul et al., 2008; Polinsky, 2008). These errors are often in areas that are particularly vulnerable for heritage speakers (i.e., gendered articles, direct object clitics, and subjunctive; Anderson, 1999a,b; Guiberson et al., 2015). Basic morphosyntactic structures are typically acquired by ages 4–5 for monolinguals (e.g., Brown, 1973; Castilla, 2008), but for bilinguals, this is a vulnerable time as children are still acquiring both of their languages and the reduced exposure to Spanish may have consequences on their grammar acquisition. Several researchers have noted that even by 6.5, many bilinguals in the US have still not acquired gendered articles expressively (Morgan et al., 2013), especially those who are more English-dominant bilinguals (Baron et al., 2018). However, there are a few studies that have shown that gendered articles are acquired by Spanish-dominant bilinguals by age 6 (Castilla-Earls et al., 2016) or when they have reached a mean utterance length of 6 (Baron et al., 2018).

Research on US heritage speakers of Spanish has mainly focused on adults (Polinsky, 2008; Montrul, 2016), with few studies focusing on child populations (e.g., Cuza and Pérez-Tattam, 2015; Baron et al., 2018, 2022; Castilla-Earls et al., 2020, 2023). Some literature suggests that heritage Spanish-speaking children do not lag behind their monolingual peers in acquiring gender up to first grade (Snyder et al., 2001; Gathercole, 2002; Kuchenbrandt, 2005). One study showed that older children, in second and third grades, appeared to have an asymmetry where they were more accurate in producing masculine agreement than feminine (Sadek, 1975). Another study showed younger children (6- to 8-year-olds) produced agreement errors, while older children (9- to 11-year-olds) produced no errors in agreement (Montrul and Potowski, 2007). Due to the variability in findings, the developmental path that heritage Spanish-speaking children take in the acquisition of specific grammatical properties continues to be unclear. Montrul et al. (2008) concluded that heritage speakers have competence in grammatical gender and agreement but that the task modality and the type of linguistic knowledge that must be established affect heritage speaker's performance due to the nature of their language acquisition experience.

To account for the observations noted previously in regard to gender sensitivity, multiple theories make predictions about the ways that listeners will respond to gender cues. One can consider grammatical gender in terms of the Competition Model, in which the utility of a cue's strength varies as a result of learning and processing (MacWhinney, 1987). Forms are initially transferred on the basis of their ability to apply to new cases. However, if this transfer leads to an error or is unnecessary, the strength of the transfer is weakened. As English does not have grammatical gender, the strength of the cue may be weakened. Four patterns emerge when considering cue strength: (1) transfer of the first language (L1) onto the second language (L2), (2) abandonment of L1 for L2, (3) merger of L1 and L2, and (4) partial attainment of separate L1 and L2 systems. The merger of L1 and L2 or partial attainment of separate L1 and L2 systems is the pattern that is most likely to affect gender processing. Early on in learning, the concept being expressed (gender sensitivity) would be more strongly associated with the form consistent with the L1 contingencies (e.g. *el gato* [the.MASC cat]) than with the form consistent with the L2 contingencies (e.g. *the cat*) (Trenkic, 2009). With more L2 experience, where gender is not marked and there is no full nominal gender system, the strength of the association in the L2 (*the cat*) is likely to increase. Another factor determining the outcome of the competition is the cognitive architecture and mechanisms involved in language processing, as well as the capacity-limited nature of working memory (MacWhinney et al., 1984). The details of the mental representation of "*el gato*" versus "*the cat*" differ across models of the mental lexicon and lexical processing. Under the Competition Model, the two competing referential forms differ in terms of their complexity, with "*el gato*" being structurally more complex than "*the cat*" due to the additional gender component. Other concurrent processes and representations may restrict the resources available for referential processing. The more demand other processes make on the limited resources, the more likely it is that they will encroach on the space needed for gender processing. This, in turn, will cause the processing of the more complex expression (*el gato*) to become increasingly unaffordable or unnecessary and leave the simpler form (*the cat*) the winner. As more cognitive resources become available with increased proficiency, there will be fewer instances in which resources are exceeded to the extent that they will completely preclude the processing of the more complex expression (*el gato*). Thus, proficiency/current language use may lead to predictable patterns of gender sensitivity. If the L1 and L2 merge, within this process, we would expect to see changing levels of gender processing and accuracy.

1.3.1 Gender asymmetry

To further expand on bilinguals' use of grammatical gender, some researchers have found an asymmetry between the use of masculine- and feminine-gendered articles. Although many researchers have not observed an asymmetry between genders, this phenomenon has been typically explained in regard to the masculine default hypothesis in which the masculine gender is considered the default or unmarked gender in Spanish (Harris, 1991), French (Hulk and Tellier, 1999), Greek (Tsimpli and Hulk, 2013), Italian (Riente, 2003), and more. Feminine gender agreement seems to be more recognizable or salient in a variety of online and offline tasks when compared to the default masculine gender (Domínguez et al., 1999; Alemán Bañón and Rothman, 2016; Beatty-Martínez and Dussias, 2019; Hur et al., 2020). Perhaps, since there is reduced input and output of the heritage language, speakers then overextend the masculine gender marking in gender agreement (Cuza and Pérez-Tattam, 2015). Within eye-tracking tasks, Spanish–English-speaking adults have been shown this gender asymmetry (to use the feminine article to facilitate processing but no facilitative process for the masculine article) (Valdés Kroff et al., 2017). Valdés Kroff et al. (2017) stated that this asymmetry may be due to an extensive overuse of the masculine gender in code-switching (as the asymmetry was noted during code-switched trials), which in turn leads speakers to ignore the gender cue during comprehension. Gender asymmetry has also been recently documented in bilingual Spanish–English-speaking children (Baron et al., 2022), where children used the feminine gender during facilitatory processing but not the masculine gender. Therefore, it seems that there is a distinction in how masculine and feminine gender are represented and processed in Spanish, stemming from distributional asymmetries (Beatty-Martínez and Dussias, 2019). In sum, it appears that the processing of grammatical gender in Spanish does indeed vary.

1.4 Grammatical gender mismatch

In addition to examining the processing of grammatical article–noun pairings, testing the processing of ungrammatical article–noun pairings, or gender mismatches, can also be informative in the study of language acquisition. Gender mismatches occur when the gender of an article and the adjacent noun are mismatched (e.g., *$el_{ART.M.SG}$ [the] $pelota_{F.SG}$ [ball]) and therefore are ungrammatical. Complementary to the study of processing with grammatical pairings, which focuses on the facilitatory effect of articles on their following noun, the study of gender mismatches can reveal sensitivity to ungrammaticality by showing inhibition. Therefore, even though it is effortful to recognize grammatical uninformative nouns, it is even harder to overcome ungrammatical article–noun pairings as it impedes comprehension for a longer time (Dahan et al., 2000; van Heugten and Shi, 2009).

Researchers who have studied the processing of gender mismatches in several gendered languages have all shown delayed noun recognition and significant inhibitory effects in both children and adults in online and offline tasks (e.g., Colé and Segui, 1994; Bates et al., 1996; Jakubowicz and Faussart, 1998; Faussart et al., 1999; Jacobsen, 1999; Dahan et al., 2000; Wicha et al., 2005; Lew-Williams and Fernald, 2007; van Heugten and Shi, 2009). In several studies, 2-year-old children exposed to French showed delayed and inhibited recognition of nouns when the article and noun were mismatched compared to when they matched (e.g., van Heugten and Shi, 2009; van Heugten and Johnson, 2011). Additionally, Guillelmon and Grosjean (2001) conducted a study with English–French adult bilinguals and monolingual French speakers. Nouns were preceded by a correct, an incorrect, or a neutral gender-marked article, and participants were asked to listen to an article–adjective–noun phrase and repeat the noun as quickly as possible. Early English–French bilinguals behaved like monolinguals in their sensitivity to gender (in matched versus mismatched trials). Late bilinguals, on the other hand, did not show a matched versus mismatched effect even when controlling for the speed of response and gender-production skills. Within an acceptability judgment task, Gómez Carrero and Ogneva (2023) found that both Russian-speaking and English-speaking L2 learners of Spanish were sensitive to gender mismatches. They speculated that an "overt morphology on the noun may act as a gender cue and facilitate the detection of gender mismatches." Thus, if ungrammatical article–noun pairings significantly impede comprehension, this may demonstrate that some article–noun dependencies have been learned and can constrain possible word candidates.

1.5 The present study

The present study examines the influence of cumulative language exposure (the number of years a child has been exposed to a language) on the processing of grammatical gender in monolingual Spanish-speaking and bilingual Spanish–English-speaking children by testing grammatical and ungrammatical article–noun pairings. In the visual world paradigm, participants are presented with a visual scene, and eye movements are recorded as they hear instructions to identify or manipulate objects on a screen. Using the visual world paradigm and behavioral measures, the present study addresses the following questions in monolingual Spanish-speaking and bilingual Spanish–English-speaking children aged 5 to 10 years old: Does cumulative language exposure to English reduce the use of gender as a cue to facilitate processing in Spanish in school-aged children? If so, do children show a differential use of gender (masculine vs. feminine)?

2 Materials and methods

2.1 Participants

Thirty monolingual Spanish-speaking children from Querétaro, Mexico, and 34 bilingual Spanish–English-speaking children from Austin, Texas, were recruited. All parents and children gave informed consent/assent to participate in the study and were compensated for their participation. This study was approved by the Institutional Review Boards at the University of Texas at Austin and the University of Rhode Island. Only participants who met the following criteria were recruited for the study: (1) ages 5–10, (2) exposure to Spanish from birth, (3) no focal brain injury, severe social–emotional problems, genetic syndromes, intellectual disability, autism spectrum disorder, hearing loss, and speech or language disorders as reported by parents, and (4) normal hearing and normal/corrected vision as reported by parents. Additionally, children were included in the study if they completed the eye tracking task ($N = 13$ excluded due to difficulty tracking their eye movements). Thus, 51 participants (19 F, 24 M, 8 unreported) were included in the final analyses for this study, which included 25 children from Mexico and 26 children from Texas.

2.2 Behavioral measures

2.2.1 Bilingual input–output survey (BIOS)

To examine a child's communication abilities in the language(s) spoken, parents completed the *Bilingual Input–Output Survey* (BIOS) in person, a questionnaire subtest within the *Bilingual Spanish English Assessment* (BESA; Peña et al., 2018). Parents detailed the history of exposure to each language a child speaks at home and school environments since birth to calculate the child's age of first exposure to English and language(s) spoken at each year of the participant's life. Parents also reported the current language input and output at home and school on an hourly basis during the week and on weekends. Cumulative English exposure is defined as the number of years a child hears and speaks English from the first year they begin to be exposed to English.

2.2.2 Expressive one-word picture vocabulary test (EOWPVT)

The Expressive One-Word Picture Vocabulary Test-Fourth Edition (EOWPVT-4; Martin and Brownell, 2011) and the Expressive One-Word Picture Vocabulary Test-4: Spanish Bilingual Edition (EOWPVT-4 SBE; Martin, 2013) are norm-referenced tests of single-word expressive vocabulary that were used to provide a gross measure of cumulative vocabulary knowledge in each language that a child speaks. After signing consent, each test was administered one time to all participants. For the current study, the EOWPVT-4 and EOWPVT-4 SBE were administered as English-only and Spanish-only versions, respectively. If the participant responded in the non-target language, the examiner redirected and prompted the participant to respond in the target language. Basal and ceiling were achieved based on the directions outlined in the test manuals for each test. The ceiling rule for the EOWPVT-4 is six consecutive incorrect responses which were especially relevant for monolingual Spanish speakers as they had very little or no exposure to English. Raw scores were calculated for each language.

2.2.3 Article–noun pair naming task

Upon conclusion of the EOWPVT, participants completed a familiarization task, which included 234 images used in the eye-tracking experiment (explained in more detail in Section 2.3). Participants were instructed to name each picture in Spanish with its corresponding definite article (el$_{\text{ART.M.SG}}$ or la$_{\text{ART.F.SG}}$). This task was included to examine whether participants were able to assign the correct/target name and gender to each image. Participants were only provided with the correct response (article+noun) if they did not know the item. If they provided a non-target name (i.e., el$_{\text{ART.M.SG}}$ coche [the car] for el$_{\text{ART.M.SG}}$ carro [the car]), they were prompted to label the item again. Participants were not asked to repeat the target name after the model was given, but many did so spontaneously. Each response was recorded, and participants' production accuracy on gender-marked articles was calculated where a score of 1 was given to correctly named images (both article and noun) and a score of 0 for an incorrectly named image (article or noun) or if the participant was unable to name the image.

2.3 Eye-tracking task

2.3.1 Materials

Thirty-six familiar nouns in Spanish (18 masculine and 18 feminine) were included as targets in the experimental stimuli. Twelve filler items were included for a total of 48 stimuli. Two practice items were presented at the beginning of the task to allow participants to become accustomed to the nature of the experiment. The target location was counterbalanced such that target stimuli appeared in each quadrant on the screen the same number of times.

Selection of noun targets and phonological competitors was restricted to words with the same initial consonant–vowel. The phonological competitor in each item had a similar syllable length to the target noun (+/− 1 syllable). The other two distractors in each stimulus also began with consonants (except for /l/ as in connected speech, the /l/ in the article "el" tends to be elongated to include the initial /l/ of the target noun becoming one word rather than two distinct words). Three presentation lists were created so that across participants, each target stimulus occurred in each condition (different, same, and ungrammatical) across lists (according to a Latin square design). The two distractors in each stimulus used for the different gender and ungrammatical gender conditions stayed together for a different target stimulus in the same-gender condition across presentation lists. Additionally, items from different categories (e.g., animals, foods, furniture, transportation, musical instruments, and clothing) were not presented together with the exception of cuchillo/cuchara [knife/spoon], pavo/pato [turkey/duck], tiburón/tigre [shark/tiger], zanahoria/salsa [carrot/salsa], and canguro/caballo [kangaroo/horse], to avoid semantic competition effects (e.g., Huettig and Altmann, 2005). However, these five items were included as they were phonological competitors with the same initial consonant–vowel and were the most closely matched selections based on noun frequency. In addition, animals are high-frequency words that most 5-year-old children are familiar with, and thus, some of the distractors were animals, even if the target or phonological competitor was an animal.

Based on the Spanish Lexical Database (Espal, Duchon et al., 2013), the phonological competitors had an average lexical frequency that was higher ($M = 1.13$, $SD = 0.70$, range = 0.08–3.33) than the frequency of the target nouns ($M = 1.01$, $SD = 0.54$, range = 0.04–2.34). The phonological competitor lexical frequency between presentation lists was not significantly different ($ps > 0.12$ in pairwise t-tests). The distractors' lexical frequency was similar across the same ($ps > 0.12$), different ($ps > 0.26$), and ungrammatical ($ps > 0.26$) gender conditions between presentation lists.

One hundred and eleven colored pictures were selected from Snodgrass and Vanderwart (1980), and an additional 123 highly imageable and concrete items were also selected. Across all three presentation lists, there were a total of 234 pictures. A subset of the items had imageability, concreteness, and familiarity ratings using EsPal (81% of the feminine items, 70% of the masculine items). Imageability, concreteness, and familiarity all use a scale from 1 to 7, with 7 indicating fully imageable, concrete, or familiar. The feminine items had average ratings of 6.16, 5.88, and 6.07, and the masculine items had average ratings of 6.16, 5.99, and 5.89 on imageability, concreteness, and familiarity, respectively. Ratings across feminine and masculine items were not significantly different. Pictures were normed for naming agreement by the first author with 10 adult and 6 child heritage Spanish speakers. Pictures with above 80% agreement were used as targets, while pictures below 80% were used in filler items. The pictures were edited to fit within 462 × 334 pixels.

Participants heard the sentence "*enséñame + el/la + target noun*" [show me + the + target noun]. A bilingual male speaker of Mexican Spanish recorded each sentence. For sentences that were ungrammatical, the article was spliced from a grammatical sentence with the same initial consonant of a different target noun and was inserted into the ungrammatical sentence so that the sentence sounded natural. The articles "*el*" and "*la*" were unstressed within the sentence. "Enséñame [show me]" was used as the instruction because in Spanish, if one says "mira a [look at] + el", the "a + el" is combined to form "al" where "a + la" stays unchanged. Therefore, "enséñame", which has been used in previous gender processing studies, was selected for the instructions (e.g., Lew-Williams and Fernald, 2007, 2010; Baron et al., 2022). For example, the participants heard 'enséñame la cama' [show me the$_{\text{ART.F.SG}}$ bed$_{\text{F.SG}}$], where "*la cama*" was the target (for all three conditions described below). For the same-gender condition, "*la casa*" [the$_{\text{ART.F.SG}}$ house$_{\text{F.SG}}$] was the phonological competitor, and *la pelota* [the$_{ART.F.SG}$ ball$_{F.SG}$] and *la jirafa* [the$_{ART.F.SG}$ giraffe$_{F.SG}$] were the distractors (Figure 1A). For the different gender condition, *el carro* [the$_{ART.M.SG}$ car$_{M.SG}$] was the phonological competitor, and *el guante* [the$_{ART.M.SG}$ glove$_{M.SG}$] and *el tenedor* [the$_{ART.M.SG}$ fork$_{M.SG}$] were the distractors (Figure 1B). For the ungrammatical condition, participants heard "enséñame *el cama" [show me the$_{ART.M.SG}$ bed$_{F.SG}$], where *la cama* was still the target, *el carro* was the phonological competitor, and *el guante* and *el tenedor* were the distractors (Figure 1B). Although the competitor and distractor pictures were the same for each item in the different-gender and ungrammatical conditions, the auditory stimulus was different. The different-gender condition had a target that was grammatical (*la cama*), while the ungrammatical condition had a target that was ungrammatical (**el cama*).

2.3.2 Design and procedure

Three groups of experimental stimuli were prepared: one group with informative (different-gender) articles, one group with uninformative (same-gender) articles, and one group with incorrect (ungrammatical) articles. A target appeared only once in each presentation list, preceded by the correct (or incorrect) gender marking surrounded by distractors with the same or different gendered articles. There were three presentation lists, and each target stimulus occurred in each condition across the lists. For example, "el conejo" [the rabbit] occurred in the same-gender condition in List 1, the different-gender condition in List 2, and the ungrammatical condition in List 3. Each participant completed one presentation list.

The experiment was built using EyeLink Experiment Builder software (v2.1.1). In Texas, eye movements were recorded on an EyeLink 1000, while in Mexico, eye movements were recorded on an EyeLink Portable Duo due to its portability. SR Research produces both eye trackers. All procedures were the same across locations. The sample rate was 500 Hz. Viewing was binocular, but eye movements were recorded monocularly. Participants were tested individually in a quiet laboratory space. Stimuli were presented on a 27-in monitor, with participants seated approximately 67 cm from the monitor with chins on a chin rest. If a child could not tolerate the chin rest, a remote tracking setting was used (only available for EyeLink Portable Duo). At the beginning of the task, participants were instructed to use a mouse to click on the object on the screen that was referred to in the sentence. To begin each trial, participants looked at a validation point in the center of the computer screen before the four images were displayed. Participants could not see the images if they did not fixate on the point in the center of the screen. Once they fixated on the point, four pictures appeared on the screen, and 500 ms later, the sentence was presented auditorily. Participants needed to click on a picture to end the trial. If a participant failed to fixate on the validation point, recalibration of the eye tracker was performed before the next trial began. In each trial, fixations were recorded from the onset of the images on the screen until the participant clicked on an image. Latencies were recorded for mouse-click responses.

2.4 Grammaticality judgment task

After the eye-tracking task, 36 targets and 12 fillers from the eye-tracking task were presented in 5–10 word sentences. Two practice sentences were presented at the beginning of the task (e.g., *Monté el camello en el desierto* [I rode the camel in the desert]). Half of the sentences were presented as grammatical sentences and half as ungrammatical sentences. Participants were asked to press one of two buttons indicating if the sentence they heard was grammatical or ungrammatical. The grammaticality across the eye-tracking task and grammaticality judgment task were the same to directly compare participants' ability to identify the grammaticality of the targets offline within sentences. If a target noun was presented in the eye-tracking task as grammatical, participants heard the noun in a simple grammatical sentence within the grammaticality judgment task. There were three grammaticality judgment lists to mirror that of the eye-tracking task, as the grammaticality of the sentence presented depended on the grammaticality of the target in the eye-tracking task. Participants were asked to listen to the sentences carefully and focus on the grammaticality of the sentence. The participant's button press was recorded for accuracy and reaction time.

2.5 Eye-tracking analyses

The eye-tracking data were exported using SR Research Data Viewer software (SR Research). An interest period was set from the beginning of the article until the participant clicked on an image. A

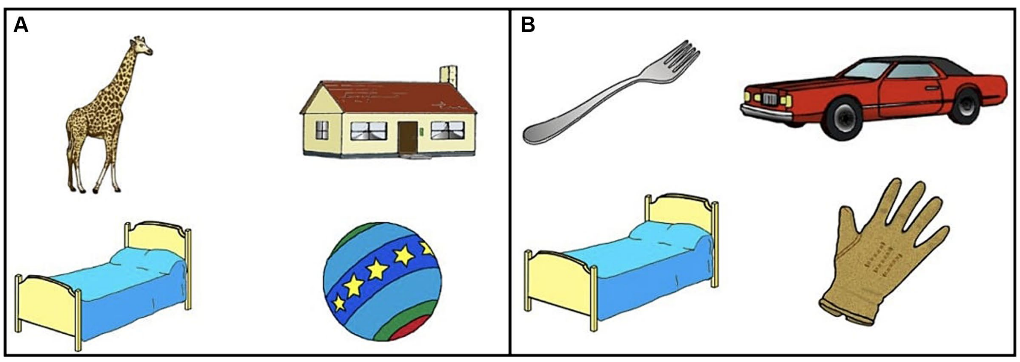


FIGURE 1
Examples of **(A)** same gender (e.g., la cama [fem.bed] as the target), **(B)** different gender (e.g., la cama [fem.bed] as the target), and ungrammatical gender (e.g., *el cama [masc.bed] as the target) displays including la cama [fem.bed] as the target noun.

Time Course (Binning) report was used to export the data. This report binned time into 10 ms bins, calculated the proportion of fixations to each image within those bins, and excluded samples that fell outside of four predefined interest areas around the images, as well as samples during blinks or saccades. Only trials where the target was correctly selected were exported for analysis (98.5% of trials for included participants).

Analyses were conducted in R (v3.5.3; R Core Team, 2021). After visual inspection of the eye tracking data, all trials with the target item *patineta* were excluded due to outlier patterns of fixations, leaving 1,759 trials. Of these, trials were excluded if the target was not correctly named in the naming task with the article and noun combination (29.2%), as the expected behavior (increased looks to the target over time) would not occur if participants did not know the target name. (Subsequent exclusion percentages in this paragraph are calculated after these exclusions.) Further data cleaning excluded trials in which participants took longer than 10 s to click on a picture (0.8%). Subsequently, trials were excluded if the participant's reaction time to click was over 2.5 *SD* of the mean click time (measured across participants on a log scale; 1.7%) or if fixation data were not present through the end of the analysis window (1,200 ms; 2.1%). The remaining 1,179 trials included in analyses were balanced between conditions: Across participants, there were between 189 and 210 trials in each of the six combinations of target gender and article gender conditions; individual participants had 3–35 trials across conditions ($M = 23$). (Note that analyses weighted each trial equally rather than weighting each participant equally, so participants with fewer trials were given less weight in analyses.)

Fixations were time-locked to the onset of the article preceding the target noun plus a 200 ms baseline (for the time it takes to plan and launch a saccade; Hallett, 1986). To increase power for the statistical analyses (a cluster permutation test, described below), the 10-ms time bins from 0 to 1,200 ms were collapsed into non-overlapping 50 ms bins; the dependent variable—indicating whether the target picture was fixated—was set to 1 if the target picture was the most-fixated interest area in at least one of the five 10-ms bins, and 0 otherwise. (This rebinning had a minimal effect on the data: Target fixations comprised 30.9% of the 10-ms bins and 34.7% of the 50-ms bins.) Finally, 50 ms time bins in which none of the interest areas were fixated were discarded (19.2% overall; between 14.6 and 22.6% in each condition).

To investigate the time course of lexical facilitation effects, looks to the target picture were analyzed with a cluster-based permutation test. These non-parametric tests were developed for the analysis of MEG/EEG data (Maris and Oostenveld, 2007; Pernet et al., 2011) but have also been applied to other time series data, including fixations in the visual world paradigm (e.g., Barr et al., 2014; see also Ito and Knoeferle, 2023) and control for multiple comparisons across time bins via permutation testing. To perform these tests in R, we used the function *clusterperm.glmer* in the *permutes* package (v2.8; Voeten, 2023). The input to this function is a generalized linear mixed-effects model, with maximal random effects, that is fit to trial- and time bin-level data (one observation for every combination of 50 ms time bin, trial, and participant). Initially, random effects are ordered by their contribution to the model, with likelihood-ratio tests used to determine whether additional random effects account for enough variance to be added to the model. Once the maximally parsimonious model is identified, the cluster test is performed using this model to test the significance of each fixed effect. The output of the test, for each factor tested, is a number of temporal "clusters" (time windows) during which the effect of that factor on fixation rates was maximal, as well as a single p-value for that factor: If that p-value is below the significance threshold (here, 0.05), then that factor is statistically significant (though the significance of the individual clusters/time bins is never directly tested; see Maris and Oostenveld, 2007 and Sassenhagen and Draschkow, 2019 for details). Several effects yielded clusters that were temporally non-contiguous but separated by only one time bin (50 ms); for reporting purposes, those clusters were combined. For transparency, the cluster mass statistic is reported for each statistically significant effect; where an effect was associated with multiple clusters, this statistic is reported for the largest cluster.

The model given as input for this analysis included a binary dependent variable (as described above, with "successes" indicating looks to the target) and four factors and their interactions. Condition (same-gender vs. different-gender vs. ungrammatical article) was Helmert-coded, such that one Condition predictor represented increased looks to the target in the different-gender condition relative to the same-gender condition, and the other Condition predictor represented increased looks to the target in the grammatical conditions (same-gender and different-gender) relative to the ungrammatical condition. Target Gender was treatment-coded (masculine = −0.5, feminine = +0.5). Cumulative English Exposure, which was originally measured on a scale of 0–10 years, was centered at the sample mean of 3.58 years and linearly scaled. Finally, age was entered as a continuous variable, centered at the sample mean of 7.65 years (range = 5.08–9.92 years). Given these contrast weights, the model intercept represents looks to the target, averaging across the levels of Condition and Gender, for participants who had 3.58 years of cumulative English exposure and were 7.65 years old. The model included these four factors and their interactions as fixed effects. It also initially included random intercepts for participants and items, as well as random slopes for all within-factor variables (for Participants: Condition, Target Gender, and their interaction; for Items: Cumulative English Exposure, Age, Condition, and their interactions). Target fixation rates for children who were 7.65 years old and had 3.58 years of cumulative English exposure, averaging across all three conditions and across both target genders, served as the baseline to which all comparisons were made.

3 Results

Table 1 shows participant means and standard deviations for age, cumulative English exposure, age of first exposure to English and Spanish, and mother education based on the Hollingshead (1975) index (a proxy for socioeconomic status) at the time of testing by geographic location. For children in Mexico, age and cumulative English exposure are not correlated ($r = 0.07$, $p = 0.75$), while for children in Texas, age and cumulative English exposure are moderately correlated ($r = 0.61$, $p = 0.001$).

Language measure (EOWPVT, grammaticality judgment, and article–noun pair naming accuracy) mean values are presented in Table 2. Grammaticality judgment accuracy is provided for grammatical and ungrammatical sentences. Grammaticality judgment sensitivity is provided as a *d'* score, and *d'* is used to compare the magnitude of discrimination ability. Here, we use *d'* to compare the magnitude of discrimination between the grammatical trials that were correctly judged

TABLE 1 Participant demographics.

	Mexico (*N* = 25)		Texas (*N* = 26)	
	Mean	*SD*	Mean	*SD*
Age (years)	7.50	1.46	7.76	1.77
Cumulative English exposure***	1.39	1.77	5.69	2.30
Age of first exposure to Spanish	0.00	0.00	0.00	0.00
English Input/Output (%)***	1.86	4.85	39.20	10.54
Mother Education[a]**	2.96	1.21	4.38	1.88

Group differences using an independent-sample t-test between the children from Mexico and Texas are denoted with $*p < 0.05$, $**p < 0.01$, and $***p < 0.001$.
[a]Measured using the following scale: 0 = not applicable or unknown, 1 = less than 7th grade, 2 = junior high school, 3 = partial high school, 4 = high school graduate, 5 = partial college, 6 = standard college or university graduation, 7 = graduate/professional training.

TABLE 2 Language measures presented in means and standard deviations.

Language measure	Mexico			Texas		
	Mean	*SD*	Range	Mean	*SD*	Range
EOWPVT English***	55.00	0.00	55–55	93.62	21.62	55–138
EOWPVT Spanish***	123.89	13.63	97–145	93.81	19.41	55–126
Grammaticality Judgment Grammatical Sentences[a]	82.82%	14.98	45.83–100	80.13%	17.49	33.33–100
Grammaticality Judgment Ungrammatical Sentences[b]	51.82%	30.32	0–100	44.55%	29.90	0–100
Grammaticality Judgment Sensitivity (*d'*)	−8.74E-08	1.57	−2.80 to 2.41	−2.18E-07	1.57	−2.77 to 2.75
Article–Noun Naming Task Accuracy (%)***	86.39%	5.54	73.93–94.02	59.66%	24.90	16.67–91.45

Group differences using an independent-sample t-test between the children from Mexico and Texas are denoted with $*p < 0.05$, $**p < 0.01$, and $***p < 0.001$.
[a]These data reflect grammatical sentences that were correctly judged to be grammatical.
[b]These data reflect ungrammatical sentences that were correctly judged to be ungrammatical.

to be grammatically correct versus the ungrammatical trials that were correctly judged to be not grammatically correct. The larger the absolute value of *d'*, the stronger the sensitivity. In both groups of children, there appears to be a low sensitivity to grammaticality within a grammaticality judgment task. Additionally, within the article–noun naming task, our sample of children from Mexico produced articles with 86.39% accuracy (*SD* = 5.54), and children from Texas produced articles with 59.66% accuracy (*SD* = 24.90).

Figure 2 shows fixations to the target object in each experimental condition. Figures 3, 4 show fixations to the target object – averaged across conditions – separately for participants with low, medium, and high residual cumulative English exposure (after regressing out the relationship with age; Figure 3) and residual age (after regressing out the relationship with cumulative English exposure; Figure 4). Both variables were treated as (unresidualized) continuous variables in analyses but were binned and residualized for ease of visualization to reflect the fact that the analysis evaluated the significance of each one while holding the other constant at the sample mean.

Both effects of the condition were significant: Participants looked at the target significantly more often in the different-gender condition than in the same-gender condition, cluster mass = 67, $p = 0.001$, an effect that was maximal from 600 to 1,000 ms. Participants also looked at the target significantly more often in the grammatical conditions than in the ungrammatical condition, cluster mass = 517, $p = 0.001$, an effect that was maximal from 600 to 1,200 ms.

The effect of Target Gender was not statistically significant, $p > 0.05$, indicating no evidence that looks to the target varied as a function of Target Gender. The extent to which participants looked at the target more in the different-gender condition than the same-gender condition was greater for items with masculine gender than items with feminine gender (a two-way interaction), maximal cluster mass = 68, $p = 0.001$, an effect that was present in three non-contiguous temporal clusters from 450 to 550, 800 to 900, and 1,100 to 1,200 ms. Target Gender did not interact with the other effect of condition (grammatical vs. ungrammatical), $p > 0.05$.

Cumulative English exposure and age each showed a significant effect but did not interact with any other factors. Relative to participants with less cumulative English exposure, participants with more cumulative English exposure looked at the target significantly less often overall, cluster mass = 219, $p = 0.001$, an effect that was maximal from 950 to 1,200 ms (see Figure 3). The effect of age was reflected in two clusters with separated time windows and opposing signs (see Figure 4). The first cluster indicated that relative to younger participants, older participants looked at the target significantly less often, with cluster mass = 34, $p = 0.001$, an effect that was maximal at 300 to 550 ms. The second cluster indicated that relative to younger participants, older participants looked at the target significantly more often, with cluster mass = 310 ms, $p = 0.001$, an effect that was maximal from 800 to 1,200 ms.

No other factors were statistically significant (all *p*s > 0.05).

4 Discussion

This study investigated the grammatical gender processing of school-aged children growing up in two different language

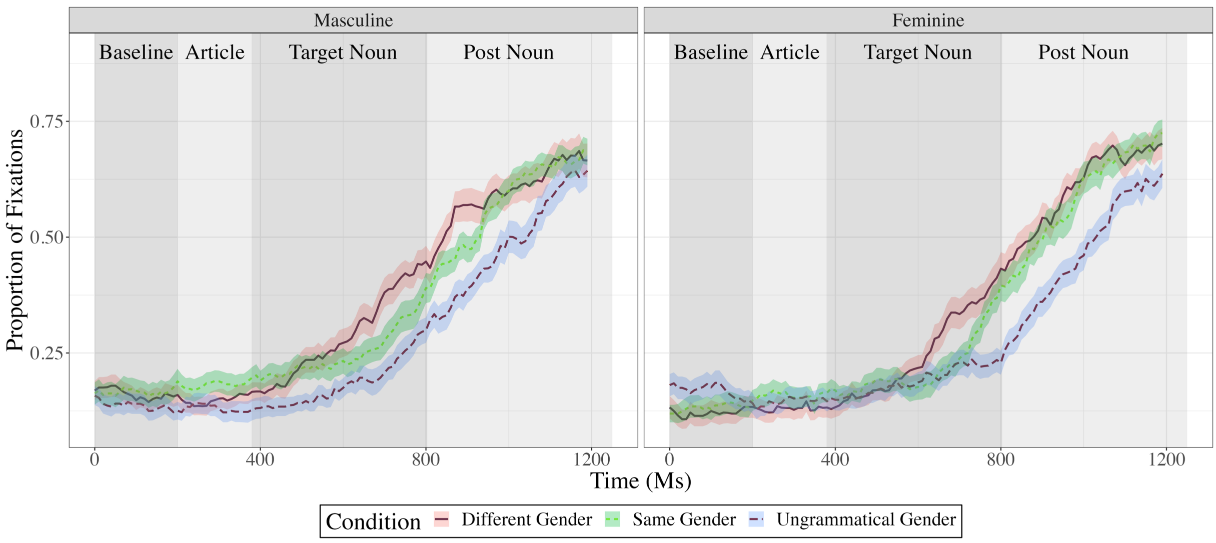


FIGURE 2
Proportion fixations to target in the different-gender (solid red), same-gender (dotted green), and ungrammatical gender (dashed blue) conditions, as a function of time in milliseconds, separately for targets with masculine gender (left panel) and feminine gender (right panel). Error ribbons represent ±1 standard error of the mean.

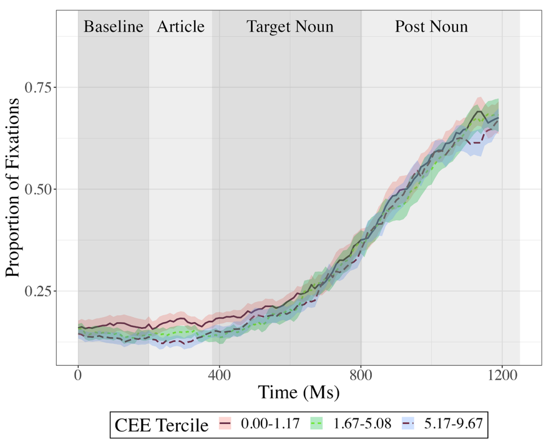


FIGURE 3
Proportion fixations to target, averaged across conditions and as a function of time in milliseconds, separately for participants based on their cumulative English exposure (CEE) residualized on age: low residual cumulative English exposure (CEE; solid red; n = 17), medium residual CEE (dotted green; n = 17), or high residual CEE (dashed blue; n = 17). Analyses treated CEE as an unresidualized continuous variable; however, participants were grouped into three bins for visualization and residualized to highlight the variance uniquely explained by CEE. Error ribbons represent ±1 standard error of the mean.

environments: in Mexico, where children were living in a community where Spanish is the dominant language, and in Texas, where children were living in a community where English is the dominant language. Previous researchers have focused on grammatical gender sensitivity to toddlers and adults; however, the literature around child heritage speakers in the US has continued to be limited. In this study, we focused on the influence of cumulative English language exposure in monolingual and heritage speakers of Spanish using grammatical and ungrammatical article–noun pairings. We addressed the research questions using a visual world paradigm in which the gendered article was informative (different gender), uninformative (same gender), or ungrammatical.

First, we examined whether cumulative language exposure to English reduced the use of the grammatical gender cue to facilitate language processing in Spanish. The results showed that all children showed lexical facilitation of informative gender marking on articles to actively anticipate an upcoming word as they looked at the target significantly more often in the different-gender condition than in the same-gender condition toward the end of the noun and during the post-noun region. Lexical processing of word recognition is affected by cumulative English exposure as children with more cumulative English exposure looked at the target noun significantly less often during the article and noun regions than children with less cumulative English exposure. Thus, their accuracy and speed to orient to the target noun were lower than those children with less cumulative exposure. Our findings show that cumulative English exposure does impact the speed and accuracy of online processing.

Additionally, relative to younger children, older children with low cumulative English exposure looked at the target noun significantly more overall during the post-noun region. Thus, older children are more accurate and look to the target nouns faster. Although monolingual children acquire grammatical gender earlier and more accurately than heritage Spanish speakers, all children in this study showed stronger lexical processing when they were older. In previous studies, children showed adult-like processing of grammatical gender but were slower than the adults, potentially due to children's slower speech processing speed and cognitive resource limitations (Lew-Williams and Fernald, 2007; Snedeker and Huang, 2015; Brouwer et al., 2017). Even though in our study we did not compare children to adults, we do see a similar difference between younger and older children.

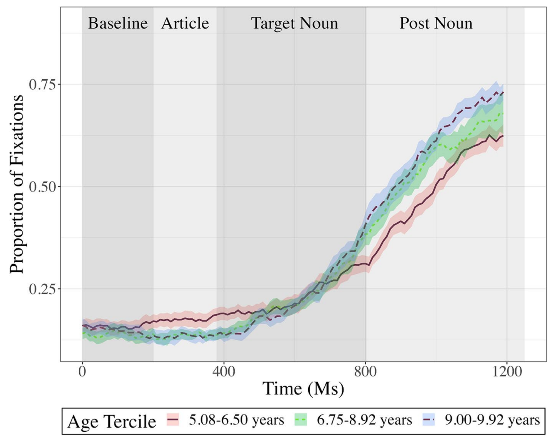


FIGURE 4
Proportion fixations to target, averaged across conditions and as a function of time in milliseconds, separately for participants based on their age residualized on cumulative English exposure: younger residual age (solid red; $n = 17$), medial residual age (dotted green; $n = 17$), or older residual age(dashed blue; $n = 17$). Analyses treated age as an unresidualized continuous variable; however, participants are grouped into three bins for visualization and residualized to highlight the variance uniquely explained by age. Error ribbons represent ±1 standard error of the mean.

In regard to overall grammaticality, all children, regardless of cumulative English exposure, looked at the target more often in the grammatical conditions than the ungrammatical conditions. Children were significantly inhibited by ungrammaticality and showed sensitivity to grammaticality within online processing but struggled during offline processing. Within the grammaticality judgment task, children did not show sensitivity to grammatical gender. This may be due to the fact that children tend to have difficulty with grammaticality judgment tasks as it significantly taxes their working memory capacity (McDonald, 2008). Children generally do not reflect on morphosyntactic structures until middle childhood (Pratt et al., 1984). Until they are older and develop an increased sensitivity to the morphosyntactic structure of sentences, children are more likely to base their judgments on semantic content and pragmatic considerations (or the plausibility of events) than on the grammaticality of the sentence.

Our findings show that children are able to use gender facilitatively but are still acquiring grammatical gender productively. Within the article–noun pair naming task, monolingual Spanish speakers were, on average, 86.39% accurate, while heritage Spanish speakers were 59.66% accurate (with a very large range of 16.67–91.45%). The path to productivity may not be a linear one, as both linguistic input and production relate to and support comprehension through early language development (Chang et al., 2006; MacDonald, 2013).

Furthermore, we investigated whether language processing differed by grammatical gender to better understand if there was a gender bias. Children with lower cumulative English exposure looked at target nouns more in the different-gender condition than the same-gender condition for masculine items more than feminine items. Thus, children with lower cumulative English exposure were facilitated more by the masculine article "*el*" than the feminine article "*la*". Other researchers have found either no asymmetry or demonstrated a gender asymmetry where they have typically found that feminine is more salient and is used to facilitate processing (Valdés Kroff et al., 2017; Baron et al., 2022). Thus, it is interesting to note that in our study, there is a gender asymmetry; however, the more salient gender appears to be masculine, and thus, children appear to be overusing masculine. Recently, Colantoni and Leroux (2024) noted that although many heritage speakers of Spanish performed at ceiling when grammatical gender was tested, a quarter of the children (aged 4;10–12;7, $M = 8;08$) still displayed a lower accuracy. They attributed this to potentially different gender grammar as some children overuse one gender, meaning they only used masculine or feminine. Within our study, we also anecdotally noticed that some children overused one gender and thus only used masculine or feminine during the article–noun pair naming task. Thus, for a subset of the children, there appears to be a "strong indication of divergence" when considering grammatical gender (Colantoni and Leroux, 2024). Other researchers have posited that heritage speakers may fail to assign any gender. The absence of gender assignment may surface as masculine morphology in Spanish, and thus, heritage speakers may appear to overuse masculine gender. This failure to assign gender may lead heritage speakers to a decreased tolerance for morphophonological irregularity (Fuchs et al., 2021).

4.1 Limitations

As noted previously, target nouns that were not named correctly during the article–noun pair naming task were excluded from the cluster-based permutation test. Given that the trial counts were already very low for several participants, the accurate naming of distractors was not considered as an additional exclusion criterion and is thus a limitation of this study.

4.2 Conclusion and future directions

In summary, school-aged children showed lexical facilitation of grammatical gender in online processing in that children looked at the target significantly more often in the different-gender condition than in the same-gender condition. Additionally, all children looked at the target significantly more often in the grammatical conditions than in the ungrammatical condition. Children with less cumulative English exposure looked at the target noun significantly more often than children with more cumulative English exposure. In regard to age, older children looked at the target noun significantly more often than younger children. Furthermore, a gender asymmetry was noted where children with less cumulative English exposure looked at the masculine target items more than feminine target items. Moreover, to the authors' knowledge, this is only the second study to use the visual world paradigm with school-aged children across the Spanish language spectrum to investigate grammatical gender. There continues to be a need to better understand Spanish language acquisition in heritage speakers of Spanish as there is a growing number of Hispanic children learning Spanish as a heritage language. Future work should continue to investigate heritage speakers across the lifespan (especially in children and adolescents) to examine the anticipation and/or facilitation of the gender cue in Spanish.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The study involving humans were approved by the University of Texas at Austin and the University of Rhode Island. The study was conducted in accordance with the local legislation and institutional requirements. Written informed consent for participation in this study was provided by the participants' legal guardians/next of kin.

Author contributions

AB: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Writing – original draft, Writing – review & editing. KC: Formal analysis, Visualization, Writing – original draft, Writing – review & editing. DK: Writing – review & editing, Formal analysis. LB: Methodology, Writing – review & editing. ZG: Methodology, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This study was financially supported by the first author's personal network through the HornRaiser crowdfunding platform at the University of Texas at Austin and the University of Rhode Island's Global Mobility Travel Grant for Faculty and Project Completion Grant.

Conflict of interest

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

Publisher's note

References

Alemán Bañón, J., and Rothman, J. (2016). The role of morphological markedness in the processing of number and gender agreement in Spanish: an event-related potential investigation. *Lang. Cogn. Neurosci.* 31, 1273–1298. doi: 10.1080/23273798.2016.1218032

Anderson, R. T. (1999a). First language loss: a case study of a bilingual child's productive skills in her first language. *Commun. Disord. Q.* 21, 4–16. doi: 10.1177/152574019902100102

Anderson, R. T. (1999b). Loss of gender agreement in L1 attrition: preliminary results. *Biling. Res. J.* 23, 389–408. doi: 10.1080/15235882.1999.10162742

Baron, A., Bedore, L. M., Peña, E. D., Lovgren-Uribe, S. D., López, A. A., and Villagran, E. (2018). Production of Spanish grammatical forms in US bilingual children. *Am. J. Speech Lang. Pathol.* 27, 975–987. doi: 10.1044/2018_AJSLP-17-0074

Baron, A., Connell, K., and Griffin, Z. M. (2022). Grammatical gender in spoken word recognition in school-age Spanish-English bilingual children. *Front. Psychol.* 13:788076. doi: 10.3389/fpsyg.2022.788076

Barr, D. J., Jackson, L., and Phillips, I. (2014). Using a voice to put a name to a face: the psycholinguistics of proper name comprehension. *J. Exp. Psychol. Gen.* 143, 404–413. doi: 10.1037/a0031813

Bates, E., Devescovi, A., Hernandez, A., and Pizzamiglio, L. (1996). Gender priming in Italian. *Percept. Psychophys.* 58, 992–1004. doi: 10.3758/BF03206827

Beatty-Martínez, A. L., and Dussias, P. E. (2019). Revisiting masculine and feminine grammatical gender in Spanish: Linguistic, psycholinguistic, and neurolinguistic evidence. *Front. Psychol.* 10:751. doi: 10.3389/fpsyg.2019.00751

Bedore, L. M., and Leonard, L. B. (2005). Verb inflections and noun phrase morphology in the spontaneous speech of Spanish-speaking children with specific language impairment. *Appl. Psycholinguist.* 26, 195–225. doi: 10.1017/S0142716405050149

Bosch, J. E., Chailleux, M., Yee, J. E., Guasti, M. T., Arosio, F., and Ayoun, D. (2022). Prediction on the basis of gender and number in mandarin-Italian bilingual children. *Stud. Biling.* 63, 243–271. doi: 10.1075/sibil.63.10bos

Bosch, J. E., and Foppolo, F. (2023). Prediction during spoken language processing in monolingual and multilingual children: investigating the role of literacy. *Linguist. Approaches Biling.* 14, 512–543. doi: 10.1075/lab.22099.bos

Brouwer, S., Sprenger, S., and Unsworth, S. (2017). Processing grammatical gender in Dutch: evidence from eye movements. *J. Exp. Child Psychol.* 159, 50–65. doi: 10.1016/j.jecp.2017.01.007

Brown, R. (1973). A first language: The early stages. Cambridge, MA: Harvard University Press.

Castilla, A. P. (2008). Developmental measures of morphosyntactic acquisition in monolingual 3-, 4-, and 5-year-old Spanish-speaking children. Unpublished doctoral dissertation, University of Toronto.

Castilla, A. P., and Pérez-Leroux, A. T. (2010). Omissions and substitutions in Spanish object clitics: developmental optionality as a property of the representational system. *Lang. Acquis.* 17, 2–25. doi: 10.1080/10489221003620904

Castilla-Earls, A., Pérez-Leroux, A. T., Martinez-Nieto, L., Restrepo, M. A., and Barr, C. (2020). Vulnerability of clitics and articles to bilingual effects in typically developing Spanish-English bilingual children. *Biling. Lang. Congn.* 23, 825–835. doi: 10.1017/S1366728919000610

Castilla-Earls, A. P., Restrepo, M. A., Perez-Leroux, A. T., Gray, S., Holmes, P., Gail, D., et al. (2016). Interactions between bilingual effects and language impairment: exploring grammatical markers in Spanish-speaking bilingual children. *Appl. Psycholinguist.* 37, 1147–1173. doi: 10.1017/S0142716415000521

Castilla-Earls, A., Ronderos, J., and Fitton, L. (2023). Spanish bilingual morphosyntactic development in bilingual children with and without developmental language disorder: articles, clitics, verbs, and the subjunctive mood. *J. Speech Lang. Hear. Res.* 66, 4678–4698. doi: 10.1044/2023_JSLHR-23-00091

Chang, F., Dell, G. S., and Bock, K. (2006). Becoming syntactic. *Psychol. Rev.* 113, 234–272. doi: 10.1037/0033-295X.113.2.234

Christou, S., Guerra, E., Coloma, C. J., Barrachina, L. A., Araya, C., Rodriguez-Ferreiro, J., et al. (2020). Real time comprehension of Spanish articles in children with developmental language disorder: empirical evidence from eye movements. *J. Commun. Disord.* 87:106027. doi: 10.1016/j.jcomdis.2020.106027

Colantoni, L., and Leroux, A. T. P. (2024). "Acoustic properties of word-final vowels and the acquisition of gender in Spanish-English heritage speakers" in Multilingual acquisition and learning: An ecosystemic view to diversity. ed. E. Babatsouli (Amsterdam: John Benjamins Publishing Company), 380–402.

Colé, P., and Segui, J. (1994). Grammatical incongruency and vocabulary types. *Mem. Cogn.* 22, 387–394. doi: 10.3758/BF03200865

Coloma, C. J., Guerra, E., De Barbieri, Z., and Helo, A. (2023). Article comprehension in monolingual Spanish-speaking children with developmental language disorder: a longitudinal eye tracking study. *Int. J. Speech Lang. Pathol.* 26, 105–117. doi: 10.1080/17549507.2023.2167235

Cornips, L., and Hulk, A. (2008). Factors of success and failure in the acquisition of grammatical gender in Dutch. *Second. Lang. Res.* 24, 267–295. doi: 10.1177/0267658308090182

Cuza, A., and Pérez-Tattam, R. (2015). Grammatical gender selection and phrasal word order in child heritage Spanish: a feature re-assembly approach. *Biling. Lang. Congn.* 19, 50–68. doi: 10.1017/S1366728914000893

Dahan, D., Swingley, D., Tanenhaus, M. K., and Magnuson, J. S. (2000). Linguistic gender and spoken-word recognition in French. *J. Mem. Lang.* 42, 465–480. doi: 10.1006/jmla.1999.2688

Bruhn De Garavito, J., and White, L. (2002). The L2 acquisition of Spanish DPs. The status of grammatical featuress. In The acquisition of Spanish morphosyntax: The L1 / L2 connection. Eds. A. T. Pérez-Leroux & J. Liceras (Dordrecht: Kluwer), 153–178.

Domínguez, A., Cuetos, F., and Segui, J. (1999). The processing of grammatical gender and number in Spanish. *J. Psycholinguist. Res.* 28, 485–498. doi: 10.1023/A:1023216326448

Duchon, A., Perea, M., Sebastián-Gallés, N., Martí, A., and Carreiras, M. (2013). EsPal: One-stop shopping for Spanish word properties. *Behav. Res. Methods* 45, 1246–1258. doi: 10.3758/s13428-013-0326-1

Faussart, C., Jakubowicz, C., and Costes, M. (1999). Gender and number processing in spoken French and Spanish. *Riv. Linguistica* 11, 75–101.

Fernald, A., Swingley, D., and Pinto, J. P. (2001). When half a word is enough: Infants can recognize spoken words using partial phonetic information. *Child Dev.* 72, 1003–1015. doi: 10.1111/1467-8624.00331

Friederici, A. D., and Jacobsen, T. (1999). Processing grammatical gender during language comprehension. *J. Psycholinguist. Res.* 28, 467–484. doi: 10.1023/A:1023264209610

Fuchs, Z., Polinsky, M., and Scontras, G. (2021). "Explaining gender: lessons from heritage Spanish" in Multifaceted multilingualism. ed. K. K. Grohmann (Amsterdam: John Benjamins Publishing Company).

Gathercole, V. C. M. (2002). "Grammatical gender in bilingual and monolingual children: A Spanish morphosyntactic distinction" in Language and literacy in bilingual children. Eds. D. K. Oller and R. E. Eilers (Clevedon, U.K: Multilingual Matters), 207–219.

Gómez Carrero, T., and Ogneva, A. (2023). Does the ending matter? Revisiting the acquisition of L2 Spanish grammatical gender by gendered and ungendered L1 adults. En A. Ariño-Bizarro, N. López-Cortés and D. Pascual (eds) En torno al lenguaje: aportaciones al estudio lingüístico. Zaragoza: Prensas de la Universidad de Zaragoza, p. 251–268.

Guiberson, M. M., Barrett, K. C., Jancosek, E. G., and Yoshinaga, I. C. (2015). Language maintenance and loss in preschool-age children of Mexican immigrants. *Commun. Disord. Q.* 28, 4–17. doi: 10.1177/15257401060280010601

Guillelmon, D., and Grosjean, F. (2001). The gender marking effect in spoken word recognition: the case of bilinguals. *Mem. Cogn.* 29, 503–511. doi: 10.3758/BF03196401

Hallett, P. (1986). Eye movements and human visual perception. In *Handbook of perception and human performance (Chap. 10)*. Eds. Boff, K. R., Kaufman, L. & Thomas, J. P. (New York: Wiley).

Harris, J. W. (1991). The exponence of gender in Spanish. *Linguist. Inq.* 22, 27–62. Available at: https://www.jstor.org/stable/4178707

Hernández-Pina, F. (1984). Teorías psicosociolingüísticas y su aplicación a la adquisición del español como lengua materna [Psychosociolinguistic theory and its application to the acquisition of Spanish as a first language]. Madrid, Spain: Siglo XXI de España.

Hollingshead, A. A. (1975). Four-factor index of social status. New Haven, CT: Yale University.

Huettig, F., and Altmann, G. T. (2005). Word meaning and the control of eye fixation: semantic competitor effects and the visual world paradigm. *Cognition* 96, B23–B32. doi: 10.1016/j.cognition.2004.10.003

Hulk, A., and Tellier, C. (1999). "Conflictual agreement in romance nominals" in Formal perspectives on romance linguistics. Amsterdam studies in the theory and history of linguistic science series. eds. J. M. Authier, B. Bullock and L. Reed (Amsterdam: Benjamins), 179–196.

Hur, E., Lopez Otero, J. C., and Sanchez, L. (2020). Gender agreement and assignment in Spanish heritage speakers: Does frequency matter? *Languages* 5:48. doi: 10.3390/languages5040048

Ito, A., and Knoeferle, P. (2023). Analysing data from the psycholinguistic visual-world paradigm: comparison of different analysis methods. *Behav. Res. Methods* 55, 3461–3493. doi: 10.3758/s13428-022-01969-3

Jacobsen, T. (1999). Effects of grammatical gender on picture and word naming: evidence from German. *J. Psycholinguist. Res.* 28, 499–514. doi: 10.1023/A:1023268310519

Jakubowicz, C., and Faussart, C. (1998). Gender agreement in the processing of spoken French. *J. Psycholinguist. Res.* 27, 597–617. doi: 10.1023/A:1023297620824

Kondo-Brown, K. (Ed.) (2006). Heritage language development. Focus on East Asian immigrants. Amsterdam: Benjamins.

Kuchenbrandt, I. (2005). Gender acquisition in bilingual Spanish. In Proceedings of the 4th international symposium on bilingualism (ISB4) (pp. 1252–1263). Somerville, MA: Cascadilla Press.

Lemmerth, N., and Hopp, H. (2019). Gender processing in simultaneous and successive bilingual children: Cross-linguistic lexical and syntactic influences. *Lang. Acquis.* 26, 21–45. doi: 10.1080/10489223.2017.1391815

Lew-Williams, C., and Fernald, A. (2007). Young children learning Spanish make rapid use of grammatical gender in spoken word recognition. *Psychol. Sci.* 18, 193–198. doi: 10.1111/j.1467-9280.2007.01871.x

Lew-Williams, C., and Fernald, A. (2010). Real-time processing of gender-marked articles by native and non-native Spanish speakers. *J. Mem. Lang.* 63, 447–464. doi: 10.1016/j.jml.2010.07.003

Lleó, C. (1998). Proto-articles in the acquisition of Spanish: interface between phonology and morphology. *Modelle der Flexion* 18, 175–195. doi: 10.1515/9783110919745.175

López-Ornat, S. (1997). "What lies in between a pre-grammatical and a grammatical representation? Evidence on nominal and verbal form-function mappings in Spanish from 1;7 to 2;1" in Contemporary perspectives on the Acquisition of Spanish, Vol 1: Developing grammars. Eds. A. T. Perez-Leroux and W. R. Glass (Sommerville, MA: Cascadilla Press), 3–20.

MacDonald, M. C. (2013). How language production shapes language form and comprehension. *Front. Psychol.* 4:40296. doi: 10.3389/fpsyg.2013.00226

MacWhinney, B. (1987). Applying the competition model to bilingualism. *Appl. Psycholinguist.* 8, 315–327. doi: 10.1017/S0142716400000357

MacWhinney, B., Bates, E., and Kliegl, R. (1984). Cue validity and sentence interpretation in English, German, and Italian. *J. Verbal Learn. Verbal Behav.* 23, 127–150. doi: 10.1016/S0022-5371(84)90093-8

Maris, E., and Oostenveld, R. (2007). Nonparametric statistical testing of EEG-and MEG-data. *J. Neurosci. Methods* 164, 177–190. doi: 10.1016/j.jneumeth.2007.03.024

Mariscal, S., and Auza, A. (2017). "Typical language development of monolingual Spanish-speaking children" in Language development and disorders in Spanish-speaking children. eds. A. Auza and R. Schwartz (Cham, Switzerland: Springer International Publishing), 3–36.

Marslen-Wilson, W. D. (1987). Functional parallelism in spoken word recognition. *Cognition* 25, 71–102. doi: 10.1016/0010-0277(87)90005-9

Martin, N. A. (2013). Expressive one-word picture vocabulary test: Spanish-bilingual edition. *4th* Edn. Novato, CA: Academic Therapy Publications.

Martin, N. A., and Brownell, R. (2011). Expressive one-word picture vocabulary test. *4th* Edn. Novato, CA: Academic Therapy Publications.

McDonald, J. L. (2008). Grammaticality judgments in children: the role of age, working memory and phonological ability. *J. Child Lang.* 35, 247–268. doi: 10.1017/S0305000907008367

Montrul, S. (2016). The acquisition of heritage languages. Cambridge: Cambridge University Press.

Montrul, S., Foote, R., and Perpiñán, S. (2008). Gender agreement in adult second language learners and Spanish heritage speakers: the effects of age and context of acquisition. *Lang. Learn.* 58, 503–553. doi: 10.1111/j.1467-9922.2008.00449.x

Montrul, S., and Potowski, K. (2007). Command of gender agreement in school-age Spanish-English bilingual children. *Int. J. Biling.* 11, 301–328. doi: 10.1177/13670069070110030301

Morgan, G., Restrepo, M. A., and Auza, A. (2013). Comparison of Spanish morphology in monolingual and Spanish–English bilingual children with and without language impairment. *Biling. Lang. Congn.* 16, 578–596. doi: 10.1017/S1366728912000697

Peña, E. D., Gutiérrez-Clellen, V. F., Iglesias, A., Goldstein, B. A., and Bedore, L. M. (2018). BESA: Bilingual English-Spanish assessment. Baltimore, MD: Brooks Publishing.

Perez-Pereira, M. (1991). The acquisition of gender: what Spanish children tell us. *J. Child Lang.* 18, 571–590. doi: 10.1017/S0305000900011259

Pernet, C. R., Chauveau, N., Gaspar, C., and Rousselet, G. A. (2011). LIMO EEG: a toolbox for hierarchical LInear MOdeling of ElectroEncephaloGraphic data. *Comput. Intell. Neurosci.* 2011:831409. doi: 10.1155/2011/831409

Polinsky, M. (2008). Gender under incomplete acquisition: heritage speakers' knowledge of noun categorization. *Heritage Lang. J.* 6, 40–71. doi: 10.46538/hlj.6.1.3

Pratt, C., Tunmer, W. E., and Bowey, J. A. (1984). Children capacity to correct grammatical violations in sentences. *J. Child Lang.* 11, 129–141. doi: 10.1017/S0305000900005626

R Core Team (2021). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.

Riente, L. (2003). Ladies first: the pivotal role of gender in the Italian nominal inflection system. In T. Grüter and T. Takehisa (Eds.), McGill working papers in linguistics, *17*, 1–54. Montreal: McGill University

Rodina, Y., and Westergaard, M. (2017). Grammatical gender in bilingual Norwegian–Russian acquisition: the role of input and transparency. *Biling. Lang. Congn.* 20, 197–214. doi: 10.1017/S1366728915000668

Sadek, C. S. (1975) The Acquisition of the Concept of grammatical gender in monolingual and bilingual speakers of Spanish [Conference presentation]. Annual Meeting of the American Educational Research Association, Washington, D.C., United States.

Sassenhagen, J., and Draschkow, D. (2019). Cluster-based permutation tests of MEG/EEG data do not establish significance of effect latency or location. *Psychophysiology* 56:e13335. doi: 10.1111/psyp.13335

Silva-Corvalán, C. (1994). Language contact and change: Spanish in Los Angeles. Oxford: Oxford University Press.

Snedeker, J., and Huang, Y. T. (2015). "Sentence processing" in The Cambridge handbook of child language. eds. E. Bavin and L. Naigles. *(2nd ed.)* ed (Cambridge, UK: Cambridge University Press).

Snodgrass, J. G., and Vanderwart, M. (1980). A standardized set of 260 pictures: norms for name agreement, image agreement, familiarity, and visual complexity. *J. Exp. Psychol. Hum. Learn. Mem.* 6:174. doi: 10.1037//0278-7393.6.2.174

Snyder, W., Senghas, A., and Inman, K. (2001). Agreement morphology and the acquisition of noun-drop in Spanish. *Lang. Acquis.* 9, 157–173. doi: 10.1207/S15327817LA0902_02

Soler, R. (1984). "Acquisicion y utilizacion del articulo" in Estudios sobre psicologia del lenguaje infantil. ed. M. Siguan (Madrid: Piramide).

Teschner, R. V., and Russell, W. M. (1984). The gender patterns of Spanish nouns: an inverse dictionary-based analysis. *Hispanic Linguist.* 1, 115–132.

Trenkic, D. (2009). Accounting for patterns of article omissions and substitutions in second language production. In Pilar García MayoM. del and R. Hawkins, eds., Second language acquisition of articles: Empirical findings and theoretical implications, 115–143. Amsterdam: John Benjamins Publishing Company

Tsimpli, I. M., and Hulk, A. (2013). Grammatical gender and the notion of default: insights from language acquisition. *Lingua* 137, 128–144. doi: 10.1016/j.lingua.2013.09.001

Unsworth, S. (2013). Assessing the role of current and cumulative exposure in simultaneous bilingual acquisition: the case of Dutch gender. *Biling. Lang. Congn.* 16, 86–110. doi: 10.1017/S1366728912000284

Unsworth, S., Argyri, F., Cornips, L., Hulk, A., Sorace, A., and Tsimpli, I. (2014). The role of age of onset and input in early child bilingualism in Greek and Dutch. *Appl. Psycholinguist.* 35, 765–805. doi: 10.1017/S0142716412000574

Valdés Kroff, J. R., Dussias, P. E., Gerfen, C., Perrotti, L., and Bajo, M. T. (2017). Experience with code-switching modulates the use of grammatical gender during sentence processing. *Linguist. Approaches Biling.* 7, 163–198. doi: 10.1075/lab.15010.val

van Heugten, M., and Johnson, E. K. (2011). Gender-marked determiners help Dutch learners' word recognition when gender information itself does not. *J. Child Lang.* 38, 87–100. doi: 10.1017/S0305000909990146

van Heugten, M., and Shi, R. (2009). French-learning toddlers use gender information on determiners during word recognition. *Dev. Sci.* 12, 419–425. doi: 10.1111/j.1467-7687.2008.00788.x

Voeten, C. C. (2023). Permutes: permutation tests for time series data. R package version 2.8. Available at: https://cran.r-project.org/package=permutes.

Wicha, N. Y. Y., Orozco-Figueroa, A., Reyes, I., Hernández, A., Gavaldón de Barreto, L., and Bates, E. (2005). When zebras become painted donkeys: grammatical gender and semantic priming interact during picture integration in a spoken Spanish sentence. *Lang. Cognit. Processes* 20, 553–587. doi: 10.1080/01690960444000241

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 14 August 2024
DOI 10.3389/fpsyg.2024.1330494

Check for updates

OPEN ACCESS

EDITED BY
Monika M. Polczynska,
University of California, Los Angeles,
United States

REVIEWED BY
Unn Røyneland,
University of Oslo, Norway
Yining Wang,
Macquarie University, Australia

*CORRESPONDENCE
Kamil Malarski
✉ kamil.malarski@amu.edu.pl

RECEIVED 30 October 2023
ACCEPTED 18 June 2024
PUBLISHED 14 August 2024

CITATION
Malarski K, Castle C, Awedyk W, Wrembel M and Jensen IN (2024) Orientation towards the vernacular and style-shifting as language behaviours in speech of first-generation Polish migrant communities speaking Norwegian in Norway.
Front. Psychol. 15:1330494.
doi: 10.3389/fpsyg.2024.1330494

Orientation towards the vernacular and style-shifting as language behaviours in speech of first-generation Polish migrant communities speaking Norwegian in Norway

Kamil Malarski[1]*, Chloe Castle[2], Witosław Awedyk[3], Magdalena Wrembel[4] and Isabel Nadine Jensen[2]

[1]Department of Sociolinguistics and Discourse Studies, Faculty of English, Adam Mickiewicz University in Poznań (AMU), Poznań, Poland, [2]Department of Language and Culture, Arctic University in Tromsø (UiT), Tromsø, Norway, [3]Institute of Linguistics, Faculty of Humanities, University of Szczecin (US), Szczecin, Poland, [4]Department of Contemporary English Language, Faculty of English, Adam Mickiewicz University in Poznań (AMU), Poznań, Poland

This study describes the patterns of dialect use among L3 Norwegian speakers born in Poland who have migrated to Norway. We collected the data in the form of sociolinguistic interviews recorded in Tromsø and Oslo, two different dialect regions, in order to examine potential differences in acquisition of two dissimilar dialects in Norwegian by L3 speakers. The analyses focus on dialectal and accentual variation in their speech, and whether frequency of dialect use is dependent on selected sociocultural factors. We have found that some speakers, especially those scoring high for overall dialect use, also display style-shifting, i.e. they use dialect features from the region more frequently in unscripted speech as opposed to in more formal speech styles elicited through reading tasks or the wordlist reading tasks. This demonstrates that language learners are capable of developing sensitivity towards the vernacular form in an L3. Moreover, it shows that first-generation migrant communities in fact may be capable of developing their L2/L3/L4 language competencies in a similar way to L1 speakers, including at the level of sociolinguistic variation.

KEYWORDS

L3 dialect acquisition, L3 Norwegian, style-shifting, vernacular, multilingualism

1 Introduction

Dialect acquisition and dialectal variation in a foreign language has not often been discussed in investigations of multilingual acquisition. In many descriptions, the topic has been neglected because a learner (a non-L1 user of a language) is expected to speak using standard forms which they have been taught from the language learning material (Husby, 2019; Balasubramanian, 2022; Milojičić, 2023). The choice and the use of the dialect is, however, vitally important for the social meanings which language communicates, i.e. it matters whether one uses the standard or the vernacular forms. A lack of attention to the topic may also be due to the popular misunderstanding that since second- or third-language structures are processed differently on a neural, cognitive and socio-cultural level, one will never attain the pragmatic competence which would enable them to fully participate in the social and cultural contexts which are constructed through the language learnt later in their lives.

In order to address the identified research gap, this article describes the acquisition and sociolinguistics of dialect use among Polish speakers of Norwegian, who were born in Poland but reside permanently in Norway. Specifically, we investigate which dialectal features are acquired first, and which last, in Norwegian spoken as an L2 and L3 by our informants, as well as to what extent these patterns reflect style-shifting (in terms of speech style or modality).

Polish people are the largest migrant group in Norway. As of 2023, there were 107,442 Polish immigrants living in Norway, and 16,583 Norwegian-born people with Polish parents (Statistics Norway, 2023a, b). Polish people relocate to Norway often for economic purposes. They mostly settle in cities, hence we have selected two urban dialects to describe how they may be acquired by this particular L3 community. Their perspective in dialect acquisition in Norwegian may be especially relevant in the light of the differences of the dialect landscapes of Poland and Norway (see Garbacz, 2014). In the former, dialects have considerably lower status in society and standard forms are encouraged. It is worth testing, therefore, whether this L3 community may be willing to use dialectal forms during the interviews.

In this article, we first explain the dialectal variation present in Norway. We then briefly explore the concept of a language 'standard' and its role in acquisition research. This is followed by a section on dialect acquisition in the L1 and Ln, and speaker orientation towards the vernacular. We then discuss dialectal variation and style shifting, before moving to the sociolinguistic situation for migrant communities in Norway. Subsequently, we outline our research questions and methodology, including information on participants, equipment, recordings, and procedure and analysis. Thereafter, we discuss results in terms of general dialect use, linguistic and sociocultural predictors, style-shifting, a dialect feature hierarchy, and individual variability. We describe these in light of our research questions, followed by limitations and conclusions.

2 Dialectal variation in Norway

Norwegian is a language characterised by high dialectal variation, the origins of which date back to the Old Norse period when, at the turn of the second millennium, national and regional variations among the North Germanic tribes began to emerge (Torp, 1998: 34ff). In the broadest sense, we could differentiate between four larger dialectal areas which are Western Norwegian – Vestnorsk, Eastern Norwegian – Østnorsk, The Trøndelag dialect – Trøndersk, and Northern Norwegian – Nordnorsk (Kristoffersen, 2000). The traditional division of Norwegian dialects may be perceived as somewhat arbitrary, as it does not account for many linguistic phenomena underlying the presence of an individual dialect feature in a given area (*cf.* Helleland and Papazian, 2005; Mæhlum and Røyneland, 2023: 27ff). Consequently, alternative classifications, such as that of Sandøy (1985) where he identifies as many as 12 dialect groups, have been under debate among Scandinavian dialectologists (*cf.* Mæhlum and Røyneland, 2023: 27ff). In other words, there is much more variation within each region. In addition to the spoken regional varieties, there are two *written* standards of Norwegian, Bokmål and Nynorsk. Bokmål is the primary language of the majority of Norwegian school children, whereas Nynorsk is the primary language of 11.6% of Norwegian school children (Statistics Norway, 2022). In both cases, children are taught the other non-primary standard as a second language form from grade eight.

It is difficult to identify a *spoken* 'standard' for Norwegian. It has sometimes been associated with the variety used in Oslo (capital; e.g. Kristoffersen, 2000), yet Johnsen (2015) shows in their metaanalysis that this may not be entirely accurate. In addition, the use of dialects is an important part of Norwegian culture. Røyneland (2017) describes how dialects work as an index of one's background and identity. They describe how changing one's dialect or mixing dialects (*knote*) is perceived as negative in the eyes of many Norwegians, as the dialects should rather remain "pure" and unchanged (Røyneland, 2017: 95–96). Speaking a non-Oslo dialect is in many contexts viewed as more Norwegian. This is mirrored in findings reported by the author where Norwegians tend to assess boys with foreign appearance as less *Norwegian* when they use the Oslo dialect compared to when they use other dialects (Røyneland, 2017: 101). The closest variety associated with the standard is the Oslo dialect, or, broadly, the South-East Norwegian dialect (Johnsen, 2015). This similarity between the Oslo dialect and what may be conceived of as a spoken standard 'standardtalemål' (Mæhlum, 2009; Sandøy, 2009) in Norwegian is an important consideration in the current study, wherein we assess acquisition of the Oslo dialect and the Tromsø dialect by L3 Norwegian speakers.

Following the traditional approach to the mapping of the Norwegian dialects, regional varieties can be identified by the set of the so-called primary and secondary distinctive features. The opposition between high and low tone, together with *tjukk* 'l' (retroflex flap [ɽ]), retroflexion and *jamvekt* ('even stress'; a prosodic feature that originates from Old Norse and affects the stress patterns of two-syllable words, particularly verbs), constitute the set of primary distinctive features between the four groups of Norwegian dialects: Eastern Norwegian, Western Norwegian, Trøndelag Norwegian, and Northern Norwegian (Kristoffersen, 2000; Mæhlum and Røyneland, 2023). Following this taxonomy, East Norwegian and Trøndelag Norwegian are classified as low-tone dialects while West Norwegian and North Norwegian are defined as high-tone dialects (Kinn and Kulbrandstad, 2023). As the above set does not suffice to account for the dialect variation of Norwegian, one may also resort to some other phonological dialect features such as palatalization, mono- or diphthongization or initial word stress in words of foreign origin. Furthermore, individual dialects can also be identified by a selection of morphophonemic features, among which the most salient are: personal pronouns (in particular first- person singular and plural, third-person feminine singular, as well as second- and third-person plural), the negative form *ikke* (variable with *ikkje*), the definite ending of feminine nouns, vowel change in the present tense forms of strong verbs and, dative endings in the noun paradigm (*cf.* Mæhlum and Røyneland, 2023). While some of these morphophonemic features do not concern the two dialects selected for the present study (the dative), other (personal pronouns and *ikke*/*ikkje* in particular) will affect the findings of the study conducted among the Polish L3 speakers in Oslo and Tromsø.

Some other selected dialect features in the two regions are presented in Table 1 (Tromsø) and Table 2 (Oslo). The dialect features

TABLE 1 Selected Tromsø dialect features.

Tromsø
palatalisation on /t/, /d/, /l/, /n/ in words like *vann*, *fjell*
lowering of /i/ to /e/, e.g. in *fisk*
lowering of /e/ to /æ/, e.g. in *sett*
Pronouns *dokker* / *dokkers* for *dere* / *deres* ('you' (pl) / 'yours' (pl))
Interrogatives *ka* / *kor* / *kæm*

TABLE 2 Selected Oslo dialect features.

Oslo
retroflexion for /rt./, /rl/, /rd/, /rn/, /rs/ -> [ʈ], [ɭ], [ɖ], [ɖ], [ɳ], [ʂ]
tjukk /l/ (retroflex flap [ɽ]), e.g. in *sola*
tone 1 – tone 2 phonemic distinction: ⟨à⟩ for accent 1, ⟨â⟩ for accent 2
-a ending in praeteritum and present perfect participles in the so-called -a verb class
-*a* ending in the definite forms of the (potentially) feminine nouns

described in Tables 1, 2 are also subject to variation within the regions due to both dialect levelling and more dialect contact, as described in Røyneland (2020). For instance, research on Northern Norwegian spoken varieties finds a development towards less palatalisation (Bull, 1996; Sætermo, 2011), and less lowering of vowels (Hårstad, 2010). Another relevant point is that retroflexion, as described in Table 2, is dominant in Norway, with the exception of the western regions.

3 Dialect acquisition and orientation towards the vernacular

The capacity of a person to acquire a dialect, irrespective of whether in their first or second language, is an ability related to acquiring a language as a whole (Chevrot and Ghimenton, 2018; Oschwald et al., 2018). There are only a few accounts which juxtapose a language and a dialect in this context (see Oschwald et al., 2018), stating that a dialect is a form of the 'standard' language differing in grammar and/or pronunciation features. Such a concept may be misleading, since every user of a language (L1 or not) speaks in a variety of a language, and 'standard' varieties are essentially the standardised forms of regional dialects as they have evolved, e.g. in England or in Norway (Trudgill, 2011; Johnsen, 2015). Notably, in both countries the dialects understood as the most standard come from the South-East, i.e. the regions surrounding the capital cities.

The question of the spoken 'standard' and the 'dialect' in Norway is rather complex and multi-layered. On the one hand, it seems that lay users of the language often emphasise that all dialects are equal, and hence few social situations require accommodation of one's dialect. On the other, the same ideologised language attitudes seem to be suppressive towards those L1 Norwegian speakers who would wish to adapt their dialect to the new local forms after having migrated within the country (Sætermo and Sollid, 2021). This may be seen in connection to the so-called emic and epic perspectives in perceptual dialectology (Cramer, 2018). What will be considered as the dialect (or the standard) in everyday interactions (emic perspective) may be a little dissimilar with what has been theorised as such by the researchers (etic perspective).

Also important to mention in the context of Norwegian is bidialectalism, i.e. speaking in two dialects of the same language (Tagliamonte and Molfenter, 2007; Nycz, 2015; Wu et al., 2016; Nycz, 2019). The discussions around bidialectism often make a distinction in whether code switching occurs between two regional dialects, or between one's regional dialect and the standardised variety (Trumper, 1989), and while, e.g. L1 speakers of Italian would code-switch depending on the social situation (Trumper, 1989), this is much less frequent in Norway (Nesse, 2023). In Norway, the dominating variety when it comes to the media, theatre, TV etc. has traditionally been Standardtalemål (also often understood as Urban Eastern Norwegian but this term is sometimes used in a wider context)[1], meaning that Norwegian speakers may still widely meet and acquire this variety irrespective of their own dialect which is then indicative of bidialectalism (Lundquist and Vangsnes, 2018); although it must be noted that the media has become much more inclusive in recent years towards the use of the dialect (see also Røyneland and Lanza, 2023). An important contribution to the contextualisation and social meanings conveyed through bidialectalism among L1 Norwegian speakers is van Ommeren (2016). It reports on how switching between the dialect forms and the more standard forms are a conscious socio-psychological process whereby Norwegian speakers build their social personas against language ideologies pertaining in a given community of speakers. A similar more recent study describes code switching between the Northern Norwegian and the South-Eastern Urban Norwegian by Tromsø children (Strand, 2022) showing how they style shift from the local forms into the South-Eastern forms, e.g. when playing. The use of dialect is vitally important in the discussion of the language use among the members of migrant communities in Norway because their dialects are intertwined with how they are viewed within the society overall. For instance, people representing foreign to Norway ethnicities are viewed more positively when speaking with an Oslo dialect than when speaking other dialects, e.g. Bergen or Valdres, as migrant groups are often expected to speak with an accented Norwegian or Standard Eastern Norwegian; they, however, are still not treated equally with ethnically native Norwegians who use the Oslo dialect in terms of the perceived dynamism (Røyneland and Jensen, 2020).

The research questions addressed in this article are based on the assumption that Ln speakers may in fact develop sensitivity towards the vernacular, and that the process is connected with the acquisition of pragmatics at the level of language processing and production and as the last component of the language structure (after semantics, syntax, pronunciation etc.). One study investigating this among transnational immigrants and their children finds that the first generation *does not* seem to adapt to phonological categories spoken in the region where they live, but the next generation does acquire this variation, and thus rejects their parent's idiosyncratic accentual patterns (Labov, 2014). This

1 It could also be argued that both Bokmål and Standardtalemål are terms better stuited for describing the morphosyntactic structures, while Urban Eastern Norwegian the term inclusive of lexis and phonology, too. Undoubtedly, Urban Eastern Norwegian is what people across the country would be most acquainted with.

vernacular reorganisation is instigated by a new source of social contact, namely, entering school and transitioning from primarily adult interaction to interaction with older peers (Denis et al., 2019). Labov (2001), citing Johnson and Newport (1989) suggests that this vernacular reorganisation stabilises at age 17, corresponding to the age at which the ability to acquire L1 syntactic intuitions has effectively ceased. However, for adults learning a new language, the development of sociolinguistic variation in the L2/Ln is understudied. Indeed, if living in the country where the L2 is spoken, they now have a new source(s) of social contact, which could possibly trigger vernacular reorganisation in the L2. Outside of the English-language context, it is also unknown as to whether they are learning the standard before moving towards the vernacular, or starting with both.

To the best of our knowledge, there is a small pool of studies indicating that Ln speakers *can and do* use dialect forms, and do so variably depending on sociolinguistic factors. In their study of second dialect acquisition in a second language, Gnevsheva et al. (2022) found that L2 speakers of English were more likely than L1 speakers to select Australian (as opposed to American) lexical items to label pictures after having lived in Australia. The L2 speakers of English were L1 Russian speakers who had started learning British English in their home country. Notably, this tendency to choose more Australian lexical items holds even for L2 American English D2 (second dialect) Australian English participants when compared with L1 American English D2 Australian English speakers. This suggests that L2 speakers are actually more likely than L1 speakers to be sensitive to and use different dialectal features, perhaps due to the fact that one's L1D1 serves more often as an identity marker (*cf.* Siegel, 2021 on L1D1 acquisition of Australian English). Another study investigating L2D2 acquisition is that of Drummond (2013), who shows that migrants with a Polish background in Manchester can and do acquire and use the Northern STRUT variant in their L2 English production, though they do so variably. Speakers are more likely to use this variant when they have a strong emotional relationship with a local person, or when they have particularly positive attitudes towards the region. We test comparable parameters in our Polish-born speakers from Oslo and Tromsø to see whether similar patterns may occur.

4 Style shifting

Use of the dialect is vitally important for conveying social meaning. In other words, it matters whether and when one uses standard forms and dialect forms and to what extent one switches between them, i.e. to what extent they style-shift. People typically use dialect forms in everyday interactions, i.e. while talking to family and closer friends, much more than in formal situations. The rates of divergence between the standard vs. dialect may be different for different languages and countries in Europe, which is dependent on many political, social (e.g. class structure), cultural and historical factors, e.g. in the Slavic languages landscape the standard is much more widely spoken than in Norway (Auer, 2011).

Popularisation of the sociolinguistic interview by Labov in his New York study (Labov, 1966) brought to light the extent to which variation occurs in a language across different language registers, or styles. Since then, this method has been adopted throughout sociolinguistics to capture the production of different linguistic variables and assess how differently they may be distributed in speech in different parts of the interview. What Labov suggested was that the more careful or formal the speech style is, the dialectal (or vernacular) features one produces. This has a lot to do with speech standardisation, school education and other social mechanisms allowing speakers of different languages to adapt the way they speak to different people and different social contexts, purely for the reasons of being understood more clearly and communicating more appropriately (Gooskens, 2018). The contexts closest to informal casual speech in a recorded interview are traditionally questions eliciting spontaneous answers, e.g. about one's lives, memories and childhood (Milroy and Gordon, 2003: 66; Labov, 2006: 70–71; Tagliamonte and Molfenter, 2007: 37–40).

What we are looking for in speech production among L3 speakers has already been shown for L1 speakers of Norwegian. In their extensive analysis, Lundquist et al. (2020) presented evidence showcasing the situations wherein Tromsø high school students would resort to using the standard, and when the regional forms would instead be used on morphological, syntactic, phonological and lexical levels. In a controlled environment, they recorded the students' speech in a few modes. The findings clearly show that Tromsø L1 speakers of Norwegian do style-shift, producing many more dialect features in unscripted spoken tasks compared to when reading texts, for example. Their strategies for style shifting differed, however, with reference to different parts of the grammar; namely, especially the production of syntactic structures was not subject to style-shifting as much as in other categories. Reversely, many dialectal morphological forms (e.g. *ka*, *kem*, *kor* as opposed to standard *hva*, *hvem*, *hvor*) were almost always preferred in the open speech tasks Lundquist et al. (2020).

5 Methodology

5.1 Aims and research questions

In the light of large dialect variation in Norway, as well as interesting social constraints under which the Polish migrant communities acquire the Norwegian dialects, we had phrased the following research questions before recruiting our informants:

1. How do L3 speakers acquire dialect features from the areas where they live?
2. How do they develop a sensitivity towards the dialect and do they use it differently in different speech registers or modalities (e.g. read vs. spoken)? If so, which dialect features are acquired earlier/later?
3. What are sociolinguistic predictors of dialect use? Does Norwegian proficiency or length of stay play a role?

We aimed to answer these questions by recruiting speakers in Oslo and Tromsø, two selected different dialect regions and assessing acquired dialect features with reference to variables such as the length of stay in Norway and proficiency in Norwegian.

5.2 Participants

Our informants comprised a group of 18 Polish-English-Norwegian speakers recorded in Oslo, and 18 recorded in Tromsø. The Oslo group included 16 female participants (all gender identities self-reported), and 2 male participants. The Tromsø group, on the other hand, included 13 female and 5 male participants. They all spoke Norwegian to an upper-intermediate or advanced level [as measured with the Norwegian Proficiency Test adapted from Language Trainers (2015)].

The Oslo group was quite uniform and displayed certain social characteristics. They had stayed in Norway for 10.8 years on average. They were aged between 31 and 63 (mean age = 40.05). Their average proficiency level for English was 18 / 24, and 25.2 / 28 for Norwegian. Many of them were engaged in higher-profile jobs; they ran their own companies (2), they were academics (3), teachers (7), engineers (2) etc. Many of them had strong ties with Polish culture and traditions. This is evidenced by the fact that a group of 13 people were recruited through a Polish Saturday school.

The Tromsø group seemed a little less uniform along social and economic scales. They were aged between 22 and 59. Their average length of stay in the Tromsø region was 4.9 years, and 6.5 years in Norway. The group was characterised with a very similar proficiency in Norwegian (average 25/28, as tested). The Tromsø group as a whole comprised academics (1), university students (9), artists (1), as well physical workers (7) from a more traditionally understood migrant community.

The participants were recruited through our extended circles and social media advertisements. They were remunerated for their participation. In the recruitment process, all were welcome regardless of their gender, ethnicity, religion or other social criteria. Their profiles are presented in Tables 3–6.

5.3 Data collection and analysis

This section outlines data collection, the applied procedures, as well as approaches to the subsequent analyses (Table 7). The data collection part lasted for about 45 min per person and included the sociolinguistic interview (c. between 15 and 35 min), followed by a sociodemographic questionnaire and proficiency tests in English and Norwegian. The sociolinguistic interview comprised three parts: a short interview in Polish, a short interview in English, and the main Norwegian component (see Table 8). The Norwegian part involved a text reading task of the Norwegian version of The North Wind and the Sun text (*Nordavinden og sola; spelled in Bokmål, progressive version*), followed by three semi-spontaneous tasks eliciting unscripted speech, i.e. narratives about free time activities, what they ate for breakfast, and daily routines. This was followed by a minimal pairs task involving production of word pairs which differ only in tone (e.g. *gjenta* / *jenta*), however this task was administered only to the Oslo group. Notably, most L1D1 (first dialect) Oslo speakers have the distinction, while in Tromsø many have minimal to no tonal distinction (Kristoffersen, 2000; Helleland and Papazian, 2005). Another difficulty is that the tone 2 feature in Norwegian is particularly difficult to learn for Ln speakers (e.g. Wetterlin, 2006, Haukland, 2016). The next task was a wordlist, which involved elicitation of pronunciation features such as retroflexion (Oslo), palatalization (Tromsø), vowel lowering (Tromsø), pronunciation of /r/ before /k/ (Tromsø). The material for the interview differed between Oslo and Tromsø with respect to the first wordlist, as different phonological categories were tested for these separate regions. The complete testing material which was administered to the participants can be seen in materials supplemented in the online repository.

The administration of the tasks was as follows; they were displayed on a 14 or 15-inch monitor in the form of a ppt presentation. The word tokens in the wordlist were shown one at a time; whereas in the minimal pairs task, two at a time. The participants were allowed to operate the presentation on their own, while the interviewer was sitting next to them, sometimes getting involved in a conversation in the open conversation tasks if the interviewee directed them, in whichever language the participant chose. This was usually the language the interviewer introduced themselves in - there were three interviewers, namely, an L1 Polish speaker, an L1 English speaker, and an L1 Norwegian speaker. Interviewers tried to use the language of the task component as much as possible. Once the sociolinguistic interview was finished, the participants filled in three online questionnaires which included (1) a sociodemographic questionnaire, (2) a proficiency test for English (the Cambridge Proficiency Test), and (3) a proficiency test for Norwegian [adapted from Language Trainers (2015)]. The sociodemographic questionnaire and the Norwegian Proficiency test can be found in Appendices 1, 2, respectively.

The equipment used throughout the interviews was a Marantz™ PMD 661 portable recorder and a SHURE™ SM35 overhead unidirectional microphone attached to it with an XLR cable. Speech was recorded at 44.1 kHz, 16-bit depth rate and saved to mono sounds in the wave format. The recorders were plugged into electric sockets while recording. The interviews were all recorded in quiet spaces in Oslo and Tromsø; including classrooms, a conference room and office rooms in university buildings.

The data analysis comprised the following steps. All the tasks were designed for the identification of dialectal features from the two selected regions, Oslo and Tromsø, respectively. In order to verify

TABLE 3 Social and language characteristics for the Oslo group.

	Mean	Range
Age (years)	39.7	31–53
Residence in Norway (in years)	12.6	1–32
AoO NO	24	3–35
AoO EN	10	3–35
AoO PL	0	–
NO proficiency	25.6 (91%)	20–28 (B2 – C2)
EN proficiency	17.75 (71%)	7–24 (A2 – C1)
% Norwegian friends	36	0–80

whether and to what extent the informants had acquired and used the dialect in the interview, each interview was listened to and rated by two Norwegian linguists familiar with both dialects. In each part of the interview, the given speaker was rated on a scale from 0 to 6, i.e. 0 when they used no dialect features from the regions, and 6 if they were using them in all applicable contexts. Later, for each speaker, their general dialect score (0–6) was calculated as the average of scores assigned to individual tasks. Hence, the dialect score was calculated based on dialect features encompassing lexical, phonological and morphosyntactic structures and how they were produced in the context of a sociolinguistic interview.

6 Results

6.1 General dialect use

The general dialect scores were calculated on the basis of the component scores marked in each task of the interview. Table 9 presents the descriptive results for dialect use across regions and styles. The data coming from each task represents the decreasing formality of speech style, i.e. from the wordlist where the most formal speech style was elicited to the unscripted speech task where the participants answered freely to everyday questions. What is noticeable is that on average the Oslo speakers displayed higher scores for dialect (M = 2.6) as compared to the Tromsø participants (M = 1.6). This variability may stem from the mere design of the interviews where different dialect features were tested for Oslo and Tromsø in the wordlist tasks, but also from somewhat different profiles of the two L3 dialect groups.

General dialect scores were computed for all participants (see Figure 1). Each speaker of Norwegian scored between 0 and 6 points. There is only one speaker for whom we did not record any dialect features from the area (AD4407AR). This speaker was in the Oslo group, was very proficient in Norwegian (score: 27/28) but had lived for less than 2 years in Norway. The lowest dialect scorers from Tromsø were HH4519IK and LF3524AL (0.2 / 6 and 0.2 / 6). They had lived in Norway for 4 and 15 years, respectively. Speaker LF3524AL was a little less proficient in Norwegian, and in formal testing attained 14 / 28 points. Other low-scoring participants were TK7710ER and TS8008UZ, recorded in Tromsø. They were fluent speakers but had spent a lot of time in the Oslo area before moving to Tromsø, which shaped their dialect considerably. In contrast, there is only one speaker who scored 5. They were recorded in Oslo, and had lived in Norway for almost 7 years. They used the language in a professional environment, working as an interpreter.

It is interesting to see that the L1 speakers of Norwegian from Tromsø do not always produce dialect features either (their mean dialect score was 4.5). The L1 Norwegian scores are presented in Figure 1 as a benchmark for comparison. This perhaps emphasises the general formal nature of the sociolinguistic interview, but also that some traditional dialect features are also undergoing change (e.g. palatalisation, lowering of /i/ to /e/) and are less and less adopted by the younger generation. The speakers below are ordered from the most frequent to the least frequent dialect users (Figure 1). It can be noted that most speakers are in fact capable of learning and using the dialect from the area where they live, yet to varying degrees. As many as 17 speakers scored 3 or more points on average. The 10 highest-scoring dialect users featured an overall dialect score of 4 or more. The speakers who scored on average higher than the L1D1 speakers of the Tromsø dialect were JM5321AR (Tromsø) AK7817SK (Tromsø), AK6923IC (Tromsø), LS5416LI (Oslo), AS6503AU (Oslo). One may instantly notice that the Oslo speakers scored higher in this index overall than the Tromsø speakers. This is due to the fact that the Oslo forms are closer structurally (in terms of morphosyntax) to the standard written forms than in the case of the Tromsø dialect. This is also reflected in the higher scores for more formal tasks like reading wordlists, for which tasks there usually is found less vernacular than in the more spoken-oriented tasks. We discuss this in Section 8.3

TABLE 4 Profession characteristics for the Oslo group.

	Number
Administrative	2
Teacher	2
Cleaning services	1
Air traffic control and administration	1
Medicine and medicine related	3
Warehouse worker	1
Editorial work (e.g. in a publishing house)	1
HR	1
Managerial	1
Researcher (at university)	2
Engineering (industrial, environmental)	1
Interpreting	1

TABLE 5 Social and language characteristics for the Tromsø group—L1 Polish group.

	Mean	Range
Age (years)	33	22–59
Residence in Norway	6 years 6 months	1 year 9 months - 16 years
AoO NO	22–35 yrs	22–36+ yrs.
AoO EN	11–14 yrs	3–36+ yrs
AoO PL	0–2 yrs	0–2 yrs
NO proficiency	24.7 (88%)	12–28 (A2 - C1)
EN proficiency	19.2 (77%)	6–25 (A2 - C2)
% Norwegian friends	38.5	2–95

TABLE 6 Profession characteristics for the Tromsø group.

	Number
Student	9
Tradesperson	1
Cleaner	2
Hospitality	1
Retail	1
Artist	1
Managerial	2
Researcher (at university)	1

TABLE 7 Sociolinguistic interview tasks in Oslo.

Order	Part
1	Reading passage (*Nordavinden og sola*)
2	Picture description task
3	Unscripted speech (*Fritid i Norge, 'Free time in Norway'*)
4	Unscripted speech (*Beskriv på norsk din daglige Rutine*, 'Tell us about your daily routine')
5	Unscripted speech (*Hva spiser du til frokost?*, 'What do you eat for breakfast?')
6	Minimal pairs for tone distinction
7	Wordlist

TABLE 8 Sociolinguistic interview tasks in Tromsø.

Order	Part
1	Reading passage (*Nordavinden og sola*)
2	Picture description task
3	Unscripted speech (*Fritid i Norge, 'Free time in Norway'*)
4	Unscripted speech (*Beskriv på norsk din daglige Rutine*, 'Tell us about your daily routine')
5	Unscripted speech (*Hva spiser du til frokost?*, 'What do you eat for breakfast?')
6	Wordlist

below, showing that for the higher-rate dialect users this trend is often in the opposite direction.

6.2 Linguistic and socio-cultural predictors

We tested for correlation between dialect scores and selected linguistic and socio-cultural variables in order to further explore the potential factors which play a role in dialect acquisition. Joint groups, rather than two separate groups representing each region, were entered into the analysis so that, potentially, some generalisations about dialect acquisition could be drawn. First, there was a strong positive relationship between the dialect score and the level of Norwegian proficiency at the level of significance ($r(36) = 0.65$, $p = 0.000018$). The data points are presented in Figure 2.

As far as the measured socio-cultural factors are concerned, a parallel correlation analysis was performed. This yielded significant moderate correlation between the length of stay in Norway, expressed in years, and the dialect score ($r(36) = 0.33$, $p = 0.049348$). The results indicate that the longer the residence in Norway, the more likely the speakers of migrant backgrounds are to use dialectal features in their Norwegian speech production (see Figure 3).

Another striking pattern we have found was the link between general dialect scores and the number of Norwegian L1 speakers as friends in participants' circles. There is a significant correlation of moderate strength between the dialect score and the percentage of Norwegian friends in one's circles ($r(36) = 0.64$, $p = 0.000026$; see Figure 4).

Additionally, we found that general dialect scoring correlated negatively with the speakers' age ($r(36) = -0.49$, $p = 0.002417$; see Figure 5). The younger the speaker, the more dialect they used. This result may be indicative of different waves of the Polish migration to Norway.

6.3 Style-shifting

Some speakers were found to style-shift in terms of their use of dialect features throughout the interview. For example, some speakers speak with greater fluency in the language in free speech than when they read aloud (e.g. MG4723AG, AJ4708RZ). Other speakers display the voiced retroflex flap [ɽ] (*tjukk /l/*) feature in open speech, but not when reading the wordlist (e.g. JS5412UL), and one speaker in Tromsø displayed palatalisation in words like *mann* ('man') when describing a picture and talking about their free time, but not when they read the wordlist (e.g. AK6923IC). Below, we present a closer examination of what exact patterns are found in style-shifting in both groups.

The two figures below represent style-shifting patterns for each speaker; Figure 6 for the Oslo group, and Figure 7 for the Tromsø group. The data points are the z-transformed dialect scores for each part of the interview for each speaker, which enabled us to better plot the deviations from the average dialect scores in each participants' performance in each task. The speech styles are presented from the most formal to the least formal speech style. The last three tasks (4, 5, 6) all represent unscripted speech. It was expected that with the first unscripted speech (semi-spontaneous) task, the interviewees would become more relaxed with every question about their hobbies, everyday life etc.

The results presented below may be interpreted in more than one way. First, style-shifting has usually been described in the context of variable use of the vernacular in different parts of a sociolinguistic interview; and this is what shows for many speakers. Namely, many high-frequency dialect users (e.g. AS6503AU, RB6114OR, MZ5021NE, TR6620AN in Oslo, and AK7817SK, BH7231LG in Tromsø) tend to display less dialect in more formal speech styles, as opposed to more casual styles. This trend was reversed for the second group (e.g. MS5129NN, ZS6219NN, MM5719AR in Oslo, and AK7817SK in Tromsø) for whom the fluctuations were recorded in the other direction, yet they still maintained the relatively high proportions for their dialect use. The third group displayed relatively constant rates for dialect use, and this most often coincided with very

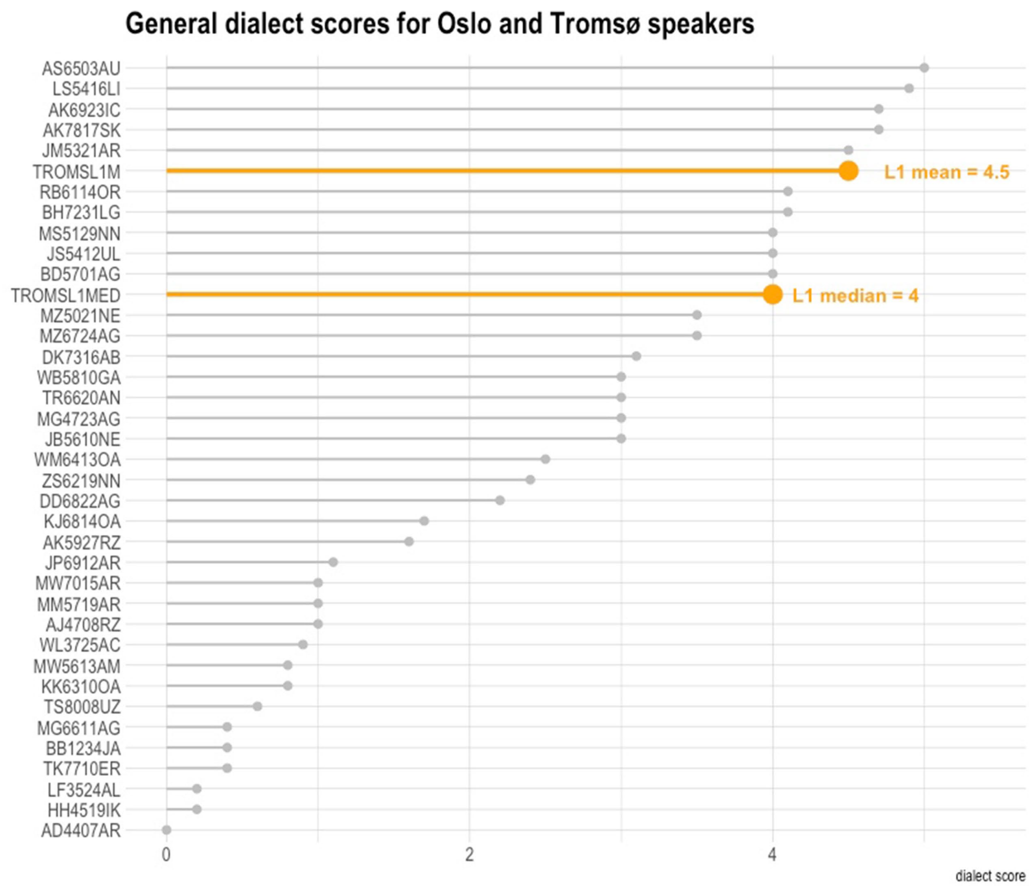


FIGURE 1
General dialect scores across all participants (L1 Norwegian scores presented as benchmark in colour).

little dialect in their repertoire (e.g. AD4407AR, DK7316AB). However, different parts of the interview may also be treated as representative of different language modalities, i.e. read vs. spoken (Figure 7).

6.4 Dialect feature hierarchy

We identified all dialect features used by our informants in Oslo and Tromsø in the free speech tasks, in order to categorise them and build a dialect feature hierarchy. Let us first look at the results from the Oslo participants. Not surprisingly, there is a clear gradation in terms of how often various features are used. First, retroflexes are featured in many speakers' L3 dialects and our data suggests that this could be the first dialect feature to be acquired by Polish learners of Norwegian when the language is learnt in the naturalistic context. Retroflexion does not appear in Polish to this degree in terms of tongue position (but there is an ongoing discussion about the phonemic - allophonic status of retroflexes in Polish, *cf.* Żygis et al., 2012), but is more frequent in Norwegian. It is not exclusively an Oslo or southeastern dialect feature, it appears in all parts of Norway, with the exception of the Western regions (Kristoffersen, 2000; Helleland and Papazian, 2005). Retroflexes appear when /r/ is followed by an alveolar or a dental consonant which gives [ɳ], [ʈ], [ɭ] etc. The second most common feature of the Oslo dialect attested in our participants' speech was the use of the tonal opposition between low tone (*lavtone*) and high tone (*høgtone*). Here, however, we have differentiated between two different situations, i.e. one in which the feature was fully acquired and one in which speakers used it variably or infrequently. It seems that many speakers have tone 1 – tone 2 differentiation, at least in some words. This is why the category of some lavtone has been included in the analysis of the collected data to account for those L3 Norwegian speakers who are able to distinguish between the two tones, but fail to use this feature consistently. The fully acquired feature, produced in all contexts applicable, is a little less frequent but is still used by more than half of the Oslo speakers (*nota bene*, we were not testing this feature among Tromsø speakers because the tonal opposition is lost or minimal in the production of many L1D1 Tromsø speakers). The next feature in the hierarchy, as used in the interviews, is the form of the feminine nouns, where they are ascribed the indefinite article *en* (for both masculine and feminine nouns *cf.* Lødrup, 2011, Rodina and Westergaard, 2015), but the feminine ending *-a* is used as a definite article, a feature which has become common even in very conservative West Oslo dialects which traditionally did not use the feminine endings in words like *jenta* ('girl'; Western, 1977; Mæhlum and Røyneland, 2023). In recent years, however, the use of -a endings in feminine nouns has been growing, also among speakers of the West Oslo dialect, most interestingly among young female residents (*cf.* Johannessen, 2008). This trend is reflected in the speech patterns of the Polish informants as many of our L3 Norwegian speakers have this form in their repertoire. Another Oslo

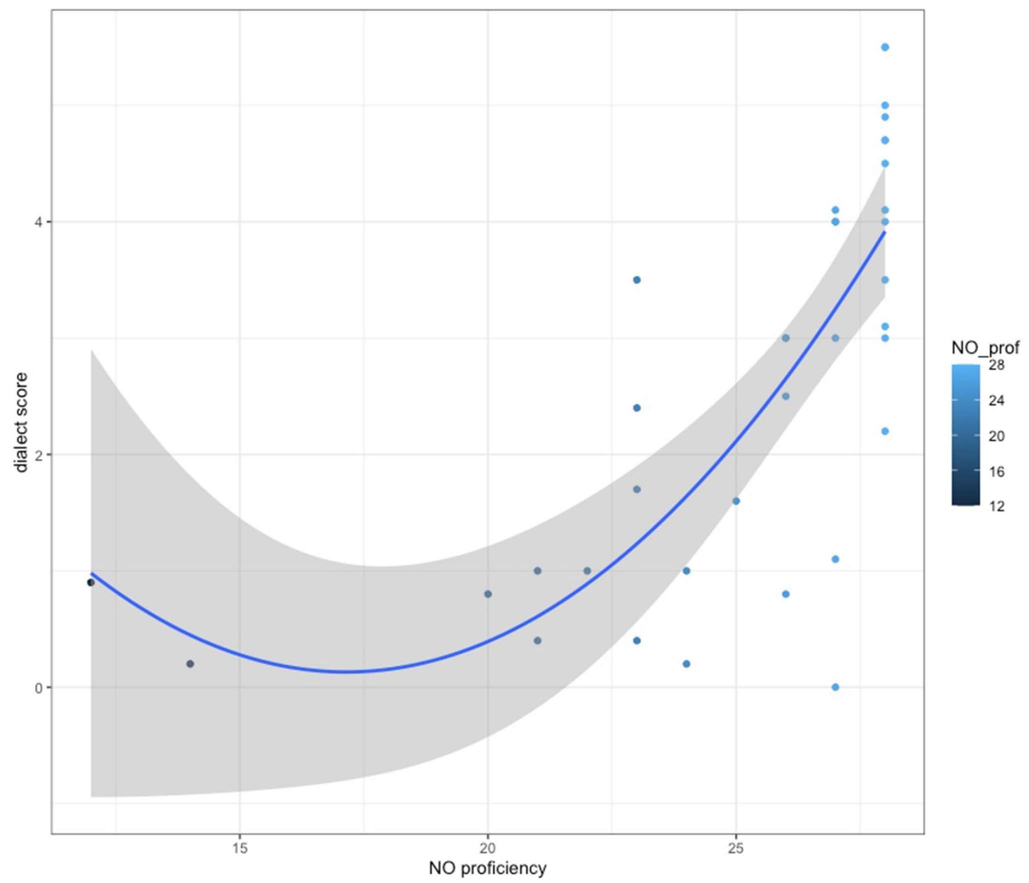


FIGURE 2
Dialect score rates against the level of Norwegian proficiency.

TABLE 9 Descriptive statistics for dialect use across regions and styles (scale: 0–6).

	Wordlist	Reading	Picture description	Unscripted speech	Overall score
Oslo					
Mean	3.8	2.1	2.3	2.2	2.6
Median	3.8	1	2	2	3
SD	1.4	1.7	1.7	1.5	1.6
Min.	1.1	0	0	0	0
Max.	5.5	5	5	4	4
Tromø					
Mean	2	2	1.9	1.6	1.6
Median	1.6	1.5	1	1	1.3
SD	1.5	1.5	2.4	1.8	1.2
Min.	0	0	0	0	0
Max.	4.6	5	6	5	4.7
Joint groups					
Mean	2.9	2.05	2.1	1.9	2.1
Median	2.7	1.3	1.5	3	2.2
SD	1.45	1.6	2.0	3.3	1.4
Min.	0	0	0	0	0
Max.	4.6	5	6	5	4.7

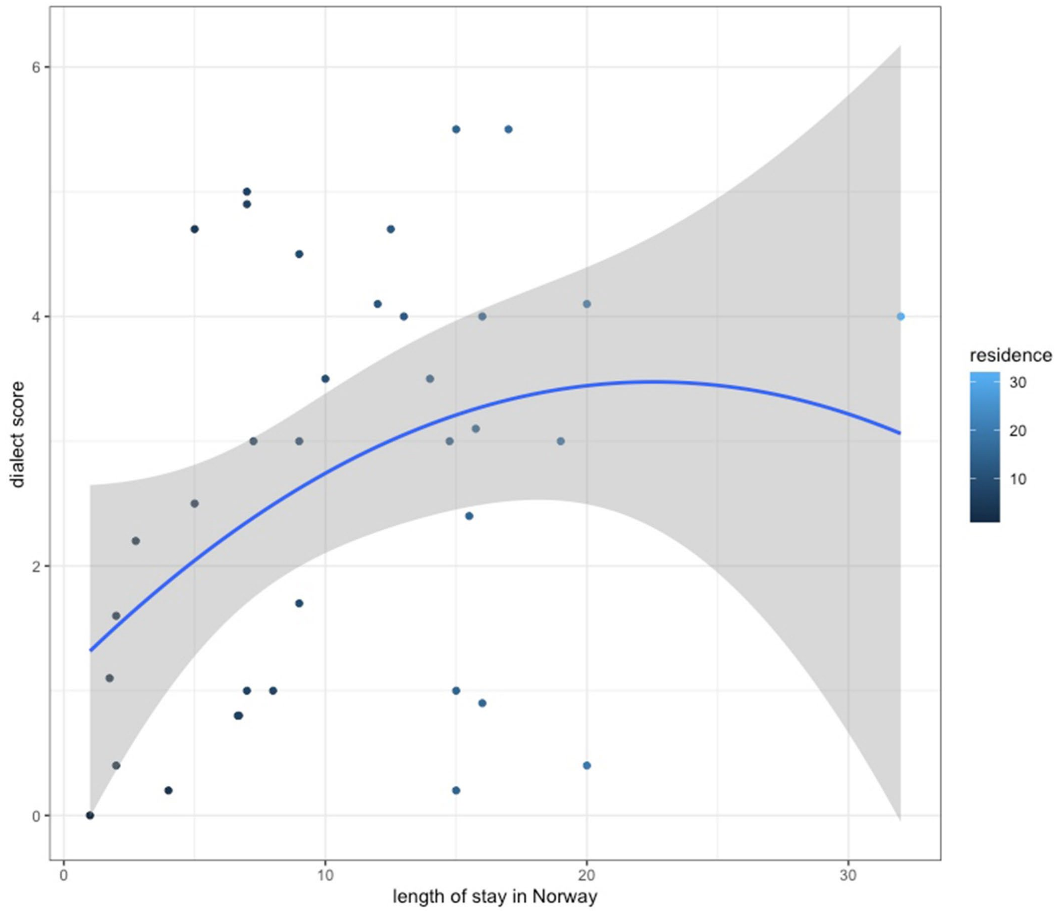


FIGURE 3
Dialect score rates against the length of stay in Norway.

dialect feature found in our participants is *tjukk /l/*, or a thick /l/, which is a retroflex flap [ɽ], used, e.g. in words like *sola*. The next Oslo-area dialect feature in the hierarchy was the replacement of long /*e*/ with a diphthong /*æi*/. Finally, there were three features which were found very infrequently in our speakers; one of these was the stress shift to the initial syllable in foreign words, a feature which is gradually disappearing among Norwegian L1 speakers and is predominantly characteristic of elder male Norwegian L1 speakers (Mæhlum and Røyneland, 2023: 61).

The hierarchy as presented in Table 10 allows us to make certain inferences about which dialectal features may be used by whom, at what frequency or in which order they are likely to be acquired. For example, if a speaker uses the voiced retroflex flap [ɽ] for /l/ (*tjukk /l/*), they most probably will have retroflexes and tonal opposition in their repertoire. The tonal opposition of Norwegian high and low tone has been acquired, though with a certain degree of inconsistency, by more than half of our Oslo speakers. This is a relatively high number for a feature which is considered especially difficult for learners of Norwegian as far as its distribution rules are concerned. It also seems that the use of feminine nouns with a definite postpositional article *-a* is a rather frequent dialect feature and, perhaps, one of the first to be acquired by learners of Norwegian living in the country. The frequency of use of these dialect features may also be indicative of how universally spread they are among L1D1 Oslo speakers. The less exposed Ln speakers are to a given feature, the less likely they are to acquire it. This is also related to the social status of each dialect feature, e.g. *tjukk /l/* (retroflex flap [ɽ]) still is a rather stereotyped dialect variant, avoided by speakers of some higher-status Oslo sociolects (Mæhlum and Røyneland, 2023: 61).

The Tromsø dialect features found in our participants are presented in Table 11. From the collected data it transpires that the Polish L3 speakers seem to score lower with regard to the use of dialect features than their counterparts from Oslo. As in the Oslo group discussed above, many Norwegian L3 speakers in Tromsø display retroflexion, a feature which is found in three main groups of Norwegian dialects. Among the most surprising findings from the Tromsø L3 speech samples, one can point to the high number of informants with the *tjukk* /l/ (retroflex flap [ɽ] sounds). Since *tjukk* /l/ is absent from the majority of Northern Norwegian dialects (Bull, 1996; Nesse and Sollid, 2010), one may infer that its occurrence in the speech samples of some Polish L3 speakers from Tromsø may be attributed to the fact that they moved to Tromsø from other parts of Norway where, most likely, the local dialect is marked by the presence of this sound. The Tromsø dialect is a variety of Norwegian referred to as *e/a mål* ('e/a variety') where infinitives have an *-e* ending, while the so-called weak feminine nouns have an *-a* ending in the indefinite form, e.g. *ei flaske/flaska* ('a bottle'). The latter dialect feature is particularly relevant to the present study since as many as

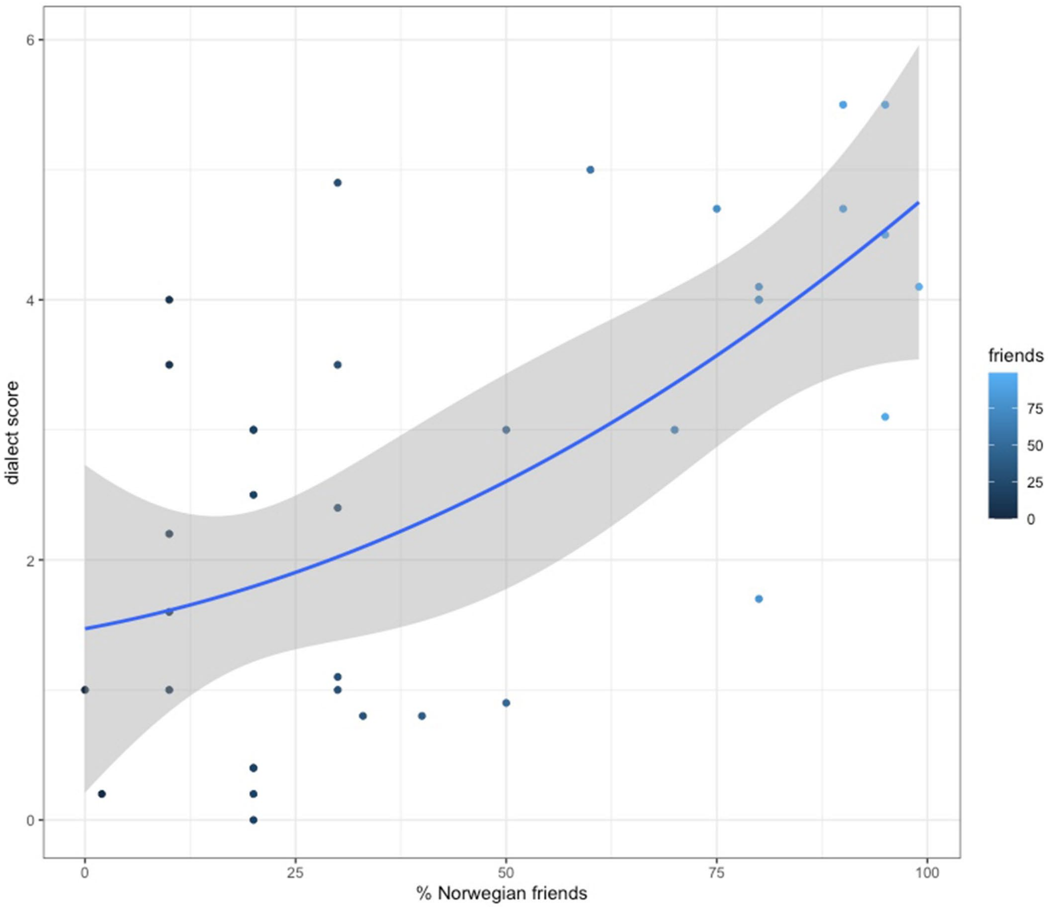


FIGURE 4
Dialect score rates against % of Norwegian friends in one's circles.

14 Polish participants from Tromsø use the indefinite forms of feminine nouns with an *-a* ending. The next feature in the hierarchy is the use of Tromsø-specific personal pronouns. Particularly the first- and third-person feminine singular pronouns score very high among our participants (*æ* 'I' and *ho* 'she'). Another feature to be discussed is the palatalisation of dentals, although we need to emphasise here that this dialect feature appears to be more characteristic of older speakers among L1 Norwegians (Bull, 1996; Jahr, 1996) which, in turn, may also explain why only three L3 speakers had it in their repertoire. Lastly, the varied and inconsistent realisation of the tone features in the L3 speech samples seem to attest to the fact that some participants may have learnt Norwegian in other regions where high-tone dialect features are not dominant. In Northern Norwegian dialects, the Old Norse /hv/ has become /kv/, or even /k/ in words like *kem* ('who'), *ka* ('what'), or *kor* ('where'). While there might be some regional variation as regards the realisation of this onset consonant cluster (*cf.* Bull, 1996; Nesse, 2008), the low frequency of this dialect feature among the Polish participants in Tromsø may be somewhat surprising. Among the dialect features which have been traditionally ascribed to the Troms region, including the urban areas of Tromsø, (*cf.* Jahr, 1996; Johannessen et al., 2009) we can mention two, which are very infrequent in the L3 speech samples. These are *lågning* (the lowering of front vowels that occurred in the Old Norse period - *cf.* Sandøy, 1985), and palatal realisation of *ikke.* This may be explained by the findings in the Nordic Dialect Corpus (Johannessen et al., 2009) where the above features can only be traced in the speech of older L1D1 speakers of the Tromsø dialect. Finally, an infrequent feature found in unscripted speech among the Tromsø participants is the dialect form in the present forms of the strong verbs (e.g. Tromsø *kjemm(er)* for *kommer*), which has been found for only two speakers.

In order to be able to compare the acquisition patterns of various dialect features, the following hierarchy of the dialect forms has been built for the L1 speakers recorded in Tromsø (Table 12). Once again, like for the L3 speakers above, the most frequently featured features come first (the left-hand columns), and are presented in the descending order. The rows, in turn, are sorted in the descending order in terms of the number of dialect features found in each speaker, i.e. the most Tromsø dialect features were recorded for GM7215OR, and the least for JRM6431UG. The dialect feature which appeared in speech of all L1 participants were retroflex sounds. The next in frequency were *høytone* (high tone), *æ* as the 1st pers. singular form, and palatalisation. It came as no surprise that the L1D1 speakers used more dialect features than the Ln speakers inhabiting the same region in Norway. However, we found no instances of the following three features in fact recorded for the L3 speakers: *nokke*, *dem / de* and the feminine form of the possessive pronoun *si*.

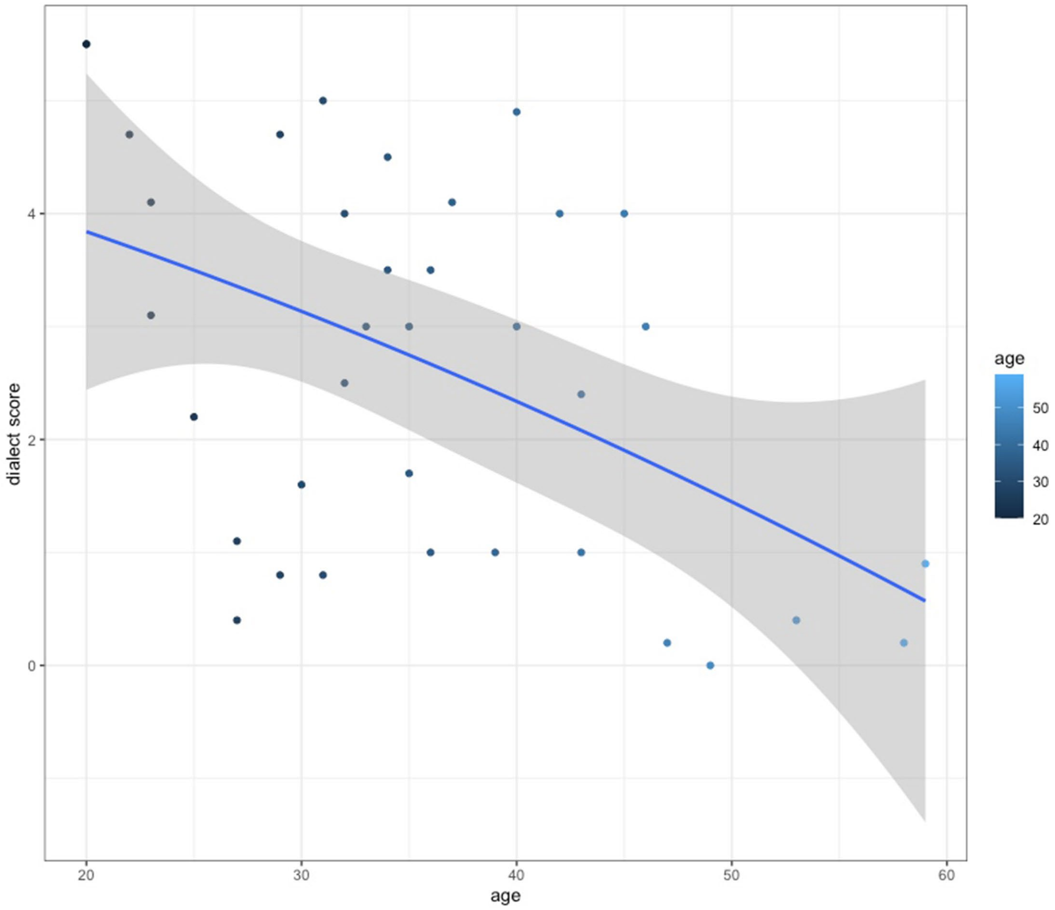


FIGURE 5
Dialect score rates against participants' age.

6.5 Individual variability

The patterns for dialect shifting, despite the generalisations drawn above, vary between different speakers. The degree of inter-speaker variation, however, is high among both L3 and L1 Norwegian speakers. Looking into several individual patterns for L3 dialect acquisition may help provide a more nuanced understanding of this process, and hence a few examples will be presented below.

A good example of a fully immersed dialect learner is speaker AK6923IC, recorded in Tromsø. At the time of the interview they had spent around 5 years in the area. They were a young person, close to finishing their university degree. They had learnt Norwegian in a naturalistic setting, speaking to their co-workers and friends. They were a fully fluent speaker, with a vast majority (90% reported in the questionnaire) of their friends being Norwegian. Their partner was Norwegian, too. Most of the time, they also used Norwegian for everyday communication. Their performance is a good example of how a high dialect user orients themselves towards the use of the vernacular. It seems that their pragmatic processing of the dialect features is similar to what would be expected from an L1 speaker from the region. In the unscripted speech tasks, we were able to detect a number of Tromsø dialect features in their performance, including the regionally used pronouns (e.g. *ho* for *hun*) or palatalisation of the word-final /l/, /n/, /k/ (e.g. in the word *hund*; Supplementary Audio 1). They also displayed considerable style-shifting, e.g. there is none or much less palatalisation in the wordlist reading (e.g. in *kann, man*; Supplementary Audio 2). They were a fully fluent Norwegian speaker (Supplementary Audio 3).

Let us now look at a Norwegian spoken by AJ4708RZ. This speaker was recorded in Oslo. They had spent a lot of time in the area, working as an engineer. They were a low dialect scoring participant. One interesting thing to be noticed about their language acquisition pattern is that they seem to be struggling with the reading tasks, reading rather slowly as if they were not a proficient language user (Supplementary Audio 4), while clearly speeding up and feeling more comfortable when speaking about their free time (Supplementary Audio 5). Their speech becomes more fluent, as if indicating how their language learning process could have been completed, namely, perhaps in a naturalistic environment at the expense of classroom instruction and the experience of reading in Norwegian. Even though they are rather fluent in the open speech tasks, their accent is retained as very L1-driven, or heavily Polish sounding. In the open speech tasks, we were able to detect only two Oslo dialect features, i.e. the diphthongisation of /e/ to /æi/, and the Oslo-like forms of the feminine nouns. In terms of their socio-economic profile, they have settled in the Oslo area and are very stable economically. They seem to be a speaker assimilated within society considering

10.3389/fpsyg.2024.1330494

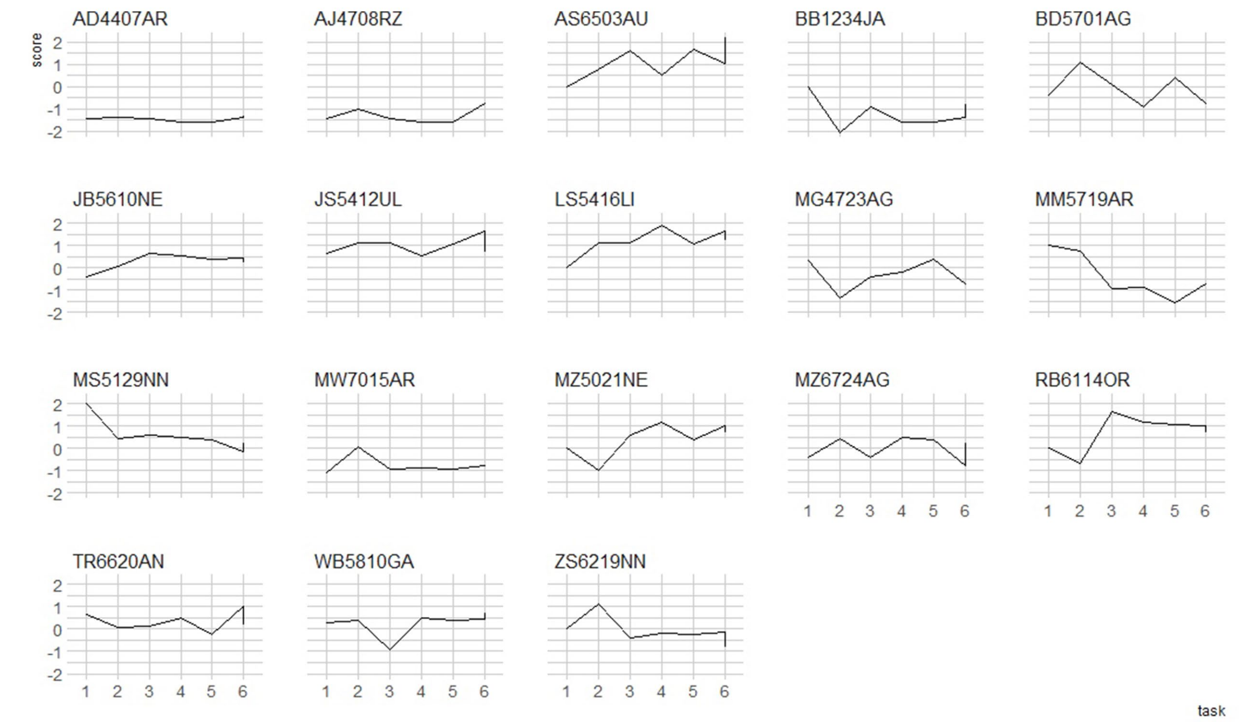


FIGURE 6
z-score plots for style shifting trajectories in the L3 Norwegian Tromsø group. 1 = wordlist, 2 = reading, 3 = picture description, 4 = unscripted (task 1), 5 = unscripted (task 2), 6 = unscripted (task 3).

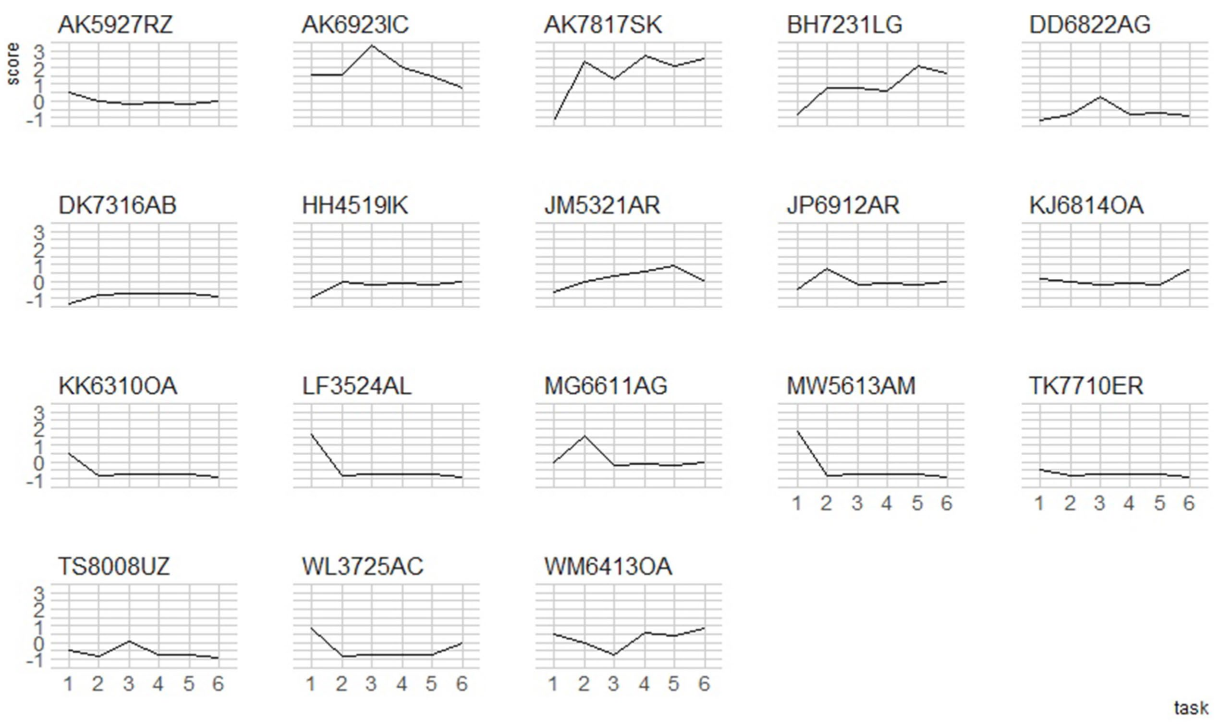


FIGURE 7
z-score plots for style shifting trajectories in the L3 Norwegian Oslo group. 1 = minimal pairs, 2 = wordlist, 3 = picture description, 4 = unscripted (task 1), 5 = unscripted (task 2).

TABLE 10 Dialect feature hierarchy for the L3 Oslo speakers.

	Retroflex	Some lavtone	F	Lavtone	*Tjukk* /l/	Diph. /e/ to /æi/	tonem 1 – tonem 2 dist.	ø	-a in past tense	Initial stress
AS6503AU	✓	✓	✓	✓	✓	✓	✓	✓	✓	
LS5416LI	✓	✓	✓	✓	✓	✓	✓	✓		
JS5412UL	✓	✓	✓	✓	✓	✓		✓		
MZ5021NE	✓	✓	✓	✓	✓		✓		✓	
RB6114OR	✓	✓	✓	✓	✓		✓		✓	
MS5129NN	✓	✓	✓	✓	✓					✓
TR6620AN	✓	✓	✓	✓	✓					
WB5810GA	✓	✓	✓	✓	✓					
ZS6219NN	✓	✓	✓	✓			✓			
JB5610NE	✓	✓	✓	✓	✓					
BD5701AG	✓	✓	✓	✓						
MG4723AG	✓	✓	✓	✓						
MZ6724AG	✓	✓			✓	✓				
MW7015AR	✓	✓								
AJ4708RZ			✓			✓				
MM5719AR	✓									
BB1234JA			✓							
AD4407AR										

their family and work status, however, they reported not having any L1 Norwegian friends in their circles which may have an impact on the acquisition of their dialect. What is also interesting is that their accent skills in English seem impressionistically comparable to their accent skills in Norwegian (Supplementary Audio 6).

In Tromsø, we recorded a subset of speakers who had spent considerable time in South-East Norway before moving to the Tromsø area. This was clearly represented in their dialect scores. Since they were permanent residents of Tromsø for a longer time, we still wanted to measure to what extent they may acquire the Tromsø dialect, or be prone to employ the strategies for dialect accommodation. In fact, we did not find many Tromsø dialect features. *Kjekla* (apocope), the use of *ikkje*, and the employment of high tone from time to time, were featured in the speech of these participants. This signals that dialect accommodation, or second dialect acquisition in an Ln, may not be a likely linguistic behaviour among similar communities living in Norway, or that it may be less likely for a person to accommodate to a dialect lower on the prestige scale (here, Tromsø) from a dialect of a higher perceived relative prestige (here, Oslo).

7 Discussion

In the first research question we wanted to investigate which varieties of the Norwegian language are spoken within the circles of Polish-born Norwegian inhabitants, comprising the largest migrant group in the country. It turned out that most of the recorded speakers did indeed use at least some dialect features from the region where they lived. Their dialect scores differed but this was predictable given that the dialect is variably used in sociolinguistic interviews because of their inherent design eliciting the vernacular to different degrees in its different parts. This result was also mirrored in the performance of the control group, i.e. L1 Tromsø speakers, for whom large variation was also recorded.

Another finding was that the level of dialect use was dependent on the language mode. Namely, for many speakers, especially the more frequent dialect users, the unscripted speech tasks where the participants were asked about their daily lives were more conducive to the use of dialect. Some speakers in Tromsø, for instance, palatalised all or some words when spontaneously answering the open questions (e.g. in the words *hund*, *mann*, *kann*) but they did not have this feature in reading, resorting to the standard nasal sound /n/. When listing down the dialect feature repertoire for each speaker (see Tables 10, 11), we focused only on the material used in the unscripted speech tasks. Building a hierarchy of dialect features, we believe, gives insight into which features are more frequently used and perhaps also acquired first in the process of dialect acquisition. This addresses the second research question: participants do seem to develop a sensitivity towards the dialect and use it differently in different speech registers.

The third research question focused on variability in dialect use along social and linguistic variables. We found a set of positively correlating factors which were the level of proficiency in Norwegian, the length of stay in the region (as opposed to length of stay in Norway overall), and the percentage of Norwegian L1 speakers present in one's circles. We analysed all the speakers, from both regions, collectively in these analyses, in order to present the potential predicting factors for dialect acquisition which would be universal, irrespective of the target dialect.

The patterns for style-shifting differ between the Oslo and the Tromsø speakers. This may be due to the design of the interview itself, but undoubtedly the status of the Oslo dialect plays a role, too, in the way that many forms used are closer to the written standard than in the case of the Tromsø dialects. One more potentially relevant factor is the clearly different sociodemographic profiles of the speakers coming from the two regions. The Tromsø participants had lived on average for a shorter time in the country and, notably, almost half of them had reported living somewhere else in Norway. This probably explains, e.g. the retroflex qualities which sounded as if acquired in a different region, as reported in a few participants. Although all the Tromsø speakers had met the recruitment criteria (having lived for longer in Norway and in the region), we did record speakers with more South-Eastern sounding accents. We did not exclude them as outliers, especially because they met the inclusion criteria for this study. Instead, we measured whether there were any Tromsø dialect features that may have developed in their repertoire. It was interesting to investigate which features would be acquired first for someone who had lived in the South-East of Norway. In a conversation outside of the recorded interview with TK7710ER, when asked if they thought they had any Tromsø features in their speech, they reported the lowering of the vowel /i/ to /e/ in their speech sometimes, and gave example of the word *fisk* ('fish'). They said this accent feature was the first feature that they noticed upon moving to the area. Overall, however, it must be concluded that these speakers have not yet accommodated their speech to the dialect spoken in the North. There are not many studies on second-dialect acquisition in a foreign language (L2D2), but Gnevsheva et al. (2022) suggested that dialect accommodation may be quicker in the second language than in the first.

Given the different status of the two dialects, one alternative interpretation of the style-shifting results could be given. Because the Oslo dialect is understood as the closest variety to the standard Norwegian speech, there would be little possible variability between the dialect and the standard forms. The fluctuations in the dialect score results then could mirror image different speech modalities, i.e. read vs. spoken. One more aspect which is different in describing style-shifting patterns between L1 and L3 speakers is also that L3 speakers have one more modality into which they may shift into, i.e. their L1 categories. For example, for those speakers who did not display retroflexion in both regions, the alternative form was not the standard, but rather their Polish-influenced pronunciations.

This data may, therefore, also point to some regularities in the acquisition of Norwegian in general. For example, many of our participants display retroflexion (which is of a different quality than in their first language). This could mean that the feature is easily perceptible by learners of Norwegian, or that it is mentioned during classroom instruction and incorporated relatively easily. Some prosodic features which are not found in speakers' L1 were also present, such as the use of *lavtone* (low tone) and *høgtone* (high tone), which we did not expect to find to such a degree.

In the light of this discussion, we interpret that dialect can be acquired by L3 speakers of Norwegian. Our data shows that L3 dialect acquisition is attainable, along with sensitivity towards the vernacular and a subconscious understanding that most speakers use the dialect to various degrees. This also signals that the process belongs to the pragmatic processing of the language. The process, however, is complex,

TABLE 11 Dialect feature hierarchy for the L3 Tromsø speakers.

	høytone	Retroflex	apocope	æ	Palatalisation	Si (gender)	dem / de	sæ	Pronouns (e.g. ho, æ)	Wh-words (kor / ka etc.)	ikkje	nokke	verb	lågning
AK7817SK	✓	✓	✓	✓	✓	✓	✓	✓	✓					✓
AK6923IC	✓	✓	✓	✓	✓				✓	✓		✓		
JM5321AR	✓	✓		✓	✓				✓				✓	✓
BH7231LG	✓	✓	✓	✓			✓			✓			✓	
DK7316AB	✓	✓	✓											
WM6413OA	✓	✓	✓											
WL3725AC			✓					✓		✓				
TK7710ER	✓	✓												
MW5613AM	✓		✓											
TS8008UZ	✓		✓											
JP6912AR		✓				✓								
DD6822AG	✓											✓		
KJ6814OA		✓									✓			
KK6310OA				✓							✓			
MG6611AG						✓	✓							
AK5927RZ	✓													
HH4519IK	✓													
LF3524AL														

TABLE 12 Dialect feature hierarchy for the L1 Tromsø speakers.

	Retroflex	høytone	æ	Palatalisation	Wh- (kor / ka etc.)	Plural forms	lågning	kjeklet	Sæ / maæ	ikkje	verb	Pronouns (e.g. ho)	-a in past tense	oppdaga
GM7215OR	✓	✓	✓	✓	✓		✓	✓	✓		✓		✓	✓
HA5809AR	✓	✓	✓	✓	✓	✓	✓	✓			✓		✓	
KF2804NN	✓	✓	✓		✓	✓	✓	✓		✓	✓		✓	
GE5012ER	✓	✓	✓	✓	✓		✓	✓	✓	✓				✓
ER6615MA	✓	✓	✓	✓		✓	✓	✓	✓	✓				
BR6119AN	✓	✓		✓	✓									
EF5520UR	✓		✓	✓	✓	✓		✓	✓			✓		
JRM6609ET	✓	✓	✓		✓			✓	✓		✓	✓		
GFL5224AI	✓	✓	✓	✓		✓	✓	✓						
KK2725EL	✓	✓	✓	✓		✓		✓				✓		
PR4819IR	✓	✓	✓	✓		✓		✓		✓				
KH5216OH	✓	✓	✓	✓			✓	✓						
JRM6431UG	✓	✓	✓					✓						

and not every L3 speaker, no matter how fluent in a foreign language (here, Norwegian), will be able to develop such a sensitivity. There were speakers (e.g. AK6923IC, AK7817SK) who use the dialect and code-switch between different tasks, as a strategy of orienting themselves towards the vernacular. What is interesting is that not all Norwegian speakers researched in Tromsø did in fact style-shift. Some used very few dialect features in their speech. This brings us once more to the notion of D1 and D2 acquisition in a first language, which is still a largely underresearched topic. Namely, some L1 speakers, just as Ln speakers, will not develop a lot of dialect, and will not be willing or capable of developing a second variety, in order to code switch or style shift in their L1. There are clearly certain cognitive and social mechanisms responsible for this (Oschwald et al., 2018), but L1 standard language ideologies and the expectations that dialects are stigmatised *per se* must also play a role (see also Auer and Røyneland, 2020).

8 Limitations

Perhaps one of the caveats behind this work is that the status of the two dialect regions compared is not exactly the same. The Tromsø dialect is a Northern dialect perceived differently in terms of prestige and of different typological structure than the Oslo dialect. Our aim initially was precisely to make use of these dissimilarities, in order to assess the potential differences in how the dialects may be learnt by Ln speakers. It could be argued, however, that the Oslo dialect is too often understood as the standard (or more standard than the Tromsø dialect), and hence the differences observed were unavoidable. The Oslo forms are structurally much closer to the Standardtalemål forms, and there are simply more forms in Tromsø which are divergent from the standard written forms. Hence the different results in dialect scoring for the two regions. Another is that we have found some complexity in participants' profiles in terms of places of residence. While most Oslo participants had lived mostly in Oslo and in the larger Oslo area, about half of the Tromsø participants had lived in other dialect regions at least for some time. They still met the inclusion criteria, having settled and lived in the Tromsø area, but the few examples reported show that speakers with these characteristics did not yet accommodate their accent to the new place. On the other hand, migration within the country has been considerable in Norway for a longer time now which undoubtedly has consequences for the development of Norwegian dialects and how they are perceived (Røyneland and Jensen, 2020; Sætermo and Sollid, 2021).

9 Conclusion

There is no one single answer to the question of how exactly the process of dialect acquisition develops in an L2 or L3. Our data coming from migrant communities speaking Norwegian in Norway points to a lot of inter- and intra-speaker variation. There are regularities, however, and therefore, this process is not entirely idiosyncratic. Especially, there are some linguistic and extra-linguistic predictors for high and low dialect use, such as the level of Norwegian proficiency, and length of residence in Norway. We demonstrate that many participants engage in style-shifting as a pragmatic strategy, using the dialect to different degrees depending on how informal the given speech act is. The implications of these findings may point to the fact that one is ready to fully assimilate with language communities in a foreign country only after understanding how sociolinguistic variation works in a foreign language; yet trying to assimilate with such communities perhaps enforces and accelerates the process of acquiring an understanding of sociolinguistic variation, too.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found at: https://github.com/kmalarski-amu/Malarski_et_al_2013.

Ethics statement

The studies involving humans were approved by Adam Mickiewicz University Ethics Committee. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

Author contributions

KM: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Resources, Visualization, Writing – original draft, Writing – review & editing. WA: Formal analysis, Investigation, Validation, Writing – original draft, Writing – review & editing. MW: Conceptualization, Funding acquisition, Resources, Supervision, Validation, Writing – review & editing. CC: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Resources, Software, Writing – original draft, Writing – review & editing. IJ: Conceptualization, Investigation, Methodology, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. The research leading to these results has received funding from the Norwegian Financial Mechanism 2014-2021 project number 2019/34/H/HS2/00495.

Acknowledgments

We would like to thank all our participants who took part in this study. We would like to thank Kamil Kaźmierski for his comments and help with data analysis and visualisation. We would like to thank Helene Jensberg for helping recruit the participants.

Conflict of interest

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

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2024.1330494/full#supplementary-material

References

Auer, P. (2011). "26 dialect vs. standard: a typology of scenarios in Europe" in *The languages and linguistics of Europe: A comprehensive guide*. eds. B. Kortmann and J. Auwera (Berlin, Boston: De Gruyter Mouton), 485–500.

Auer, P., and Røyneland, U. (2020). Modelling acquisition and use of dialectal, standard and multiethnolectal features in migratory contexts across Europe. *J. Multiling. Multicult. Dev.* 1–8, 1–8. doi: 10.1080/01434632.2020.1730385

Balasubramanian, C. (2022). "World Englishes in the EFL Classroom: The reality," in *World Englishes, Global classrooms: The future of English literary and linguistic studies*. Springer Nature. 3–15.

Bull, T. (1996). "Målet i Troms og Finnmark" in *Nordnorske dialektar*. eds. E. H. Jahr and O. Skare (Oslo: Novus Vorlag), 157–174.

Chevrot, J., and Ghimenton, A. (2018). "Bilingualism and Bidialectalism" in *The Cambridge handbook of bilingualism*. eds. HouwerA. De and L. Ortega (Cambridge: Cambridge University Press), 510–523.

Cramer, J. (2018). "The emic and the etic in perceptual dialectology," in *Language regard: Methods, variation and change*. eds. B. E. Evans, E. J. Benson and J. N. Stanford (chapter, Cambridge: Cambridge University Press), 62–79.

Denis, D., Gardner, M. H., Brook, M., and Tagliamonte, S. A. (2019). Peaks and arrowheads of vernacular reorganization. *Lang. Var. Chang.* 31, 43–67. doi: 10.1017/S095439451900005X

Drummond, R. (2013). The Manchester polish STRUT dialect Acquisition in a Second Language. *J. Engl. Linguist.* 41, 65–93. doi: 10.1177/0075424212449172

Garbacz, P. (2014). Dialekter i Norge og i Polen – forskjellig status? [Dialects in Norway and in Poland–Different status?]. *NOA norsk som andrespråk* 30, 24–39.

Gnevsheva, K., Szakay, A., and Jansen, S. (2022). Lexical preference in second dialect acquisition in a second language. *Int. J. Biling.* 26, 163–180. doi: 10.1177/13670069211036932

Gooskens, C. (2018). "Dialect intelligibility" in *The handbook of dialectology*. eds. C. Boberg, J. Nerbonne and D. Watt (Oxford: Wiley-Blackwell), 204–218.

Hårstad, S. (2010). Unge språkbrukere i gammel by: en sosiolingvistisk studie av ungdoms talemål i Trondheim, [young speakers in old city: a sociolinguistic study of teenagers spoken variety in Trondheim] 2010, 141. Available at: https://ntnuopen.ntnu.no/ntnu-xmlui/handle/11250/243705

Haukland, O. (2016). Attitudes to Norwegian-accented English among Norwegian and non-Norwegian listeners. [master's thesis]. [Oslo]: University of Oslo.

Helleland, B., and Papazian, E. (2005). Norsk talemål: Lokal og sosial variasjon [spoken Norwegian: local and social variation]. *Høyskoleforlaget.*

Husby, O. (2019). Talt Bokmål: Målspråk for andresråksstudenter? [Spoken Bokmål: Target language for second language students?]. NOA-Norsk Som andrespråk (2), Available at: https://ojs.novus.no/index.php/NOA/article/view/1710

Jahr, E. H. (1996). Dialektane i indre Troms – Bardu og Målselv. [Dialects in Indre Troms - Bardu and Målselv] in: *Nordnorske dialekta [Norwegian dialects]*. eds. E. H. Jahr and O. Skare, O. (Oslo: Novus), 180–185.

Johannessen, J. B. (2008). The pronominal psychological demonstrative in Scandinavian: Its syntax, semantics and pragmatics. *Nord. J. Linguist.* 31, 161–192. doi: 10.1017/S0332586508001923

Johannessen, J. B., Priestley, J. J., Hagen, K., Åfarli, T. A., and Vangsnes, Ø. A.. (2009). "The nordic dialect corpus–an advanced research tool," in *Proceedings of the 17th Nordic conference of computational linguistics (NODALIDA 2009)*, 73–80.

Johnsen, S. S. (2015). Dialect change in south-East Norway and the role of attitude in diffusion. *J. Socioling.* 19, 612–642. doi: 10.1111/josl.12156

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

Kinn, T., and Kulbrandstad, L. A. (2023). *Språkmønster. Innføring i det norske språksystemet. Language patterns. Introduction to the linguistic system of Norwegian*. Oslo: Universitetsforlaget.

Kristoffersen, G. (2000). *The phonology of Norwegian*. Oxford: Oxford University Press.

Labov, W. (1966). *The social stratification of English in new York City*. Washington, D.C.: Centre for Applied Linguistics.

Labov, W. (2001). *Principles of linguistic change, volume II: Social factors*. Oxford: Blackwell.

Labov, W. (2006). *The social stratification of English in new York City*. *2nd* Edn. Cambridge: Cambridge University Press.

Labov, W. (2014). "The sociophonetic orientation of the language learner" in *Advances in Sociophonetics*. eds. S. Calamai and C. Celat (Amsterdam/Philadelphia: John Benjamins), 17–28.

Language Trainers. (2015). Online Norwegian Level Test. Available at: https://www.languagetrainers.com/norwegian-level-test.php]

Lødrup, H. (2011). Hvor mange genus er det i Oslo-dialekten? ['How many genders does the Oslo dialect have?']. *Maal og Minne* 2, 120–136.

Lundquist, B., and Vangsnes, Ø. A. (2018). Language separation in Bidialectal speakers: evidence from eye tracking. *Front. Psychol.* 9:1394. doi: 10.3389/fpsyg.2018.01394

Lundquist, B., Westendorp, M., and Strand, B.-M. S. (2020). Code-switching alone cannot explain intraspeaker syntactic variability: evidence from a spoken elicitation experiment. *J. Nord. Linguist.* 43, 249–287. doi: 10.1017/S0332586520000190

Mæhlum, B. (2009). Standardtalemål? Naturligvis! En argumentasjon for eksistensen av et norsk standardtalemål [standard spoken language? Certainly! Arguing for the existence of a standard spoken language in Norway]. *Norsk Lingvistisk Tidsskrift* 27, 7–26.

Mæhlum, B., and Røyneland, U. (2023). "Det norske dialektlandskapet. Innføring i studiet av dialekter. 2 utgave, The Norwegian dialect landscape" in *An introduction to the study of dialects*. *2nd* ed (Oslo: Cappelen Damm Akademisk). Available at: https://www.akademika.no/humaniora/sprak/det-norske-dialektlandskapet/9788202746568

Milojičić, V. (2023). Deconstructing the myth of Standard German: Navigating language ideologies in the L2 German university classroom. *Fore. Lang. Ann.* 56, 453–479. doi: 10.1111/flan.12665

Milroy, L., and Gordon, M. (2003). *Sociolinguistics. Method and interpretation*. Oxford: Blackwell Publishing.

Nesse, A. (2008). *Bydialekt, Riksmål og Identitet-Sett frå Bodø [Urban dialect, riksmål and identity - seen from Bodø]*. Oslo: Novus.

Nesse, A. (2023). Norwegian. *Oxford Res. Encyclopedia of Linguistics*. doi: 10.1093/acrefore/9780199384655.013.944

Nesse, A., and Sollid, H. (2010). Nordnorske bymål i komparativt perspektiv [northern Norwegian urban dialects in a comparative perspective]. *Maal og Minne* 1, 137–158.

Nycz, J. (2015). Second dialect acquisition: a Sociophonetic perspective. *Lang Ling Compass* 9, 209–218. doi: 10.1111/lnc3.12133

Nycz, J. (2019). Media and second dialect acquisition. *Annu. Rev. Appl. Linguist.* 39, 152–160. doi: 10.1017/S0267190519000060

Oschwald, J., Schättin, A., Von Bastian, C. C., and Souza, A. S. (2018). Bidialectalism and bilingualism: Exploring therole of language similarity as a link between linguistic ability and executive control. *Front. Psychol.* 9:1997. doi: 10.3389/fpsyg.2018.01997

Rodina, Y., and Westergaard, M. (2015). Grammatical gender in Norwegian: language acquisition and language change. *J. Germanic Linguis.* 27, 145–187. doi: 10.1017/S1470542714000245

Røyneland, U. (2017). Hva ska til for å høres ut som du hører til? Forestillinger om dialektale identiteter i det senmoderne Norge. [What should you sound like to sound like you belong? Conceptions of dialectal identities in late modern Norway]. *Nordica Helsingensia* 48, 91–106.

Røyneland, U., and Jensen, B. U. (2020). Dialect acquisition and migration in Norway – questions of authenticity, belonging and legitimacy. *J. Multiling. Multicult. Dev.* 1–17. doi: 10.1080/01434632.2020.1722679

Røyneland, U., and Lanza, E. (2023). "Dialect diversity and migration: disturbances and dilemmas, perspectives from Norway" in *Language, society and the state in a changing world*. eds. S. D. Brunn and R. Kehrein (Cham: Springer), 337–355.

Røyneland, U. (2020) in *"Regional varieties in Norway revisited" in intermediate language varieties: Koinai and regional standards in Europe*. eds. M. Cerruti and S. Tsiplakou (Amsterdam: John Benjamins), 31–54.

Sætermo, M. (2011). n, l, t og d. En komparativ studie av palataler versus postalveolarer i Alta-dialekten. [Masteravhandling i nordisk språk]. Tromsø: UiT Norges arktiske universitet.

Sætermo, M., and Sollid, H. (2021). Reported language attitudes among Norwegian speaking in-migrants in Tromsø. *Acta Borealia* 38, 60–80. doi: 10.1080/08003831.2021.1911209

Sandøy, H. (1985). *Norsk dialektkunnskap [Norwegian dialectology]*. Oslo: Novus.

Sandøy, H. (2009). Standardtalemål? Ja, men...! Ein definisjon og ei drøfting av begrepet. *Norsk Lingvistisk Tidsskrift* 27, 27–48.

Siegel, J. (2021). "Identity, authenticity and dialect acquisition: the case of Australian English" in *Sociolinguistic variation and language acquisition across the lifespan*. eds. A. Ghimenton, A. Nardy and J. Chevrot (Amsterdam: John Benjamins), 277–294.

Statistics Norway. (2022). *Pupils in primary and lower secondary school.* Available at: https://www.ssb.no/en/utdanning/grunnskoler/statistikk/elevar-i-grunnskolen.

Statistics Norway. (2023a). Facts about Immigration. Available at: https://www.ssb.no/en/innvandring-og-innvandrere/faktaside/innvandring

Statistics Norway. (2023b). Norwegian for adult immigrants. Available at: https://www.ssb.no/en/utdanning/voksenopplaering/statistikk/norskopplaering-for-voksne-innvandrere

Strand, Bror-Magnus S. (2022). The roles role play plays. The form and function of bilectal codeswitching in north Norwegian pre-school children's role play. PhD UiT The Arctic University of Norway. Available at: https://munin.uit.no/bitstream/handle/10037/25836/thesis.pdf?sequence=3&isAllowed=y

Tagliamonte, S. A., and Molfenter, S. (2007). How'd you get that accent?: acquiring a second dialect of the same language. *Lang. Soc.* 36, 649–675. doi: 10.1017/S0047404507070911

Torp, A. (1998). *Nordiske språk i nordisk og germansk perspektiv [Scandinavian languages in a Scandinavian and Germanic perspective]*. Oslo: Novus.

Trudgill, P. (2011). *Sociolinguistic typology: Social determinants of linguistic complexity*. Oxford: Oxford University Press.

Trumper, J. (1989). Observations on sociolinguistic behaviour in two Italian regions. *Ital. Socioling. Tren. Iss.* 76, 31–62. doi: 10.1515/ijsl.1989.76.31/html

van Ommeren, R. (2016). Den flerstemmige språkbrukeren: En sosiolingvistisk studie av norske bidialektale. PhD dissertation, NTNU.

Western, K. (1977). *A-endinger i Oslo-mål [A-endings in Oslo dialect]. Skrifter fra talemålsundersøkelsen i Oslo (TAUS) 5*. Oslo: Novus.

Wetterlin, A. (2006). The lexical specification of Norwegian tonal word accents. [dissertation], [Konstanz]: University of Konstanz.

Wu, Y. G., Zhang, H., and Guo, T. (2016). Does speaking two dialects in daily life affect executive functions? An event-related potential study. *PLoS One* 11:e0150492. doi: 10.1371/journal.pone.0150492

Żygis, M., Pape, D. L., and Jesus, L. (2012). (non)retroflex Slavic affricates and their motivation. Evidence from Czech and polish. *J. Int. Phon. Assoc.* 42, 281–329. doi: 10.1017/S0025100312000205