CrossMark

The dynamics of spoken word recognition in bilinguals

Amy S. Desroches[1] , Deanna C. Friesen[2] , Matthew Teles[1], Chloe A. Korade[1] and Evan W. Forest[3]

[1]Department of Psychology, The University of Winnipeg, Winnipeg, Canada; [2]Faculty of Education, The University of Western Ontario, London, Canada and [3]Department of Psychology, University of Manitoba, Winnipeg, Canada

Abstract

This study examined whether bilinguals automatically activate lexical options from both of their languages when performing a picture matching task in their dominant language (L1) by using event related potentials. English–French bilinguals and English monolinguals performed a picture-spoken word matching task with three conditions: match (BEACH-"beach"), unrelated mismatch (BEACH-"tack"), and L2 onset competitor mismatch (BEACH-"plaid"; *plaid* sounds like *plage*, the French word for *beach*). Critically, bilinguals, but not monolinguals, showed reduced N400s for L2-cohort vs. unrelated mismatches. The results provide clear evidence that when bilinguals identify pictures, they automatically activate lexical options from both languages, even when expecting oral input from only their dominant language. N400 attenuation suggests bilinguals activate but do not expect L2 lexical options.

Introduction

Spoken word recognition is continuous and dynamic: as spoken words unfold, lexical alternatives compete for recognition (TRACE, McClelland & Elman, 1986; Allopenna, Magnuson & Tanenhaus, 1998). For example, as *beach* unfolds, cohorts, words with similar initial phonology, compete for recognition (e.g., *bead*). As the word continues, rhymes and other neighbors are activated (e.g., *peach, batch*). Competition suggests an interaction between top-down lexical representations and bottom-up phoneme processing. For bilinguals, competition is not restricted to one language (e.g., for an English–French bilingual, *beach* may also activate second language competitors like *bien*; Marian & Spivey, 2003a; Spivey & Marian, 1999). The question of the basis of between-language connections is, therefore, important: do bilinguals have separate language-specific lexical representations, or do they have lexical representations that encode phonology from both languages, and a language selection mechanism that allows words to be identified correctly from the target language? We used event-related potentials (ERPs) during a picture/spoken word matching task to investigate the dynamics of cross-language connections in bilinguals.

Temporally sensitive measures like eyetracking and ERPs have been used to investigate spoken word processing within languages. Eyetracking has revealed that phonologically similar words compete for recognition, marked by more looks to items that are phonologically related to targets (e.g., cohorts and rhymes) vs. unrelated distractors (Allopenna et al., 1998; Desroches, Joanisse & Robertson, 2006). ERPs have further characterized the role of phonological similarity, with two components being relevant to the current study. The Phonological Mapping Negativity (PMN) is a negative going deflection that occurs in the 200–400ms range, over frontal/central sites (see Lewendon, Mortimore & Egan, 2020). The PMN reflects bottom-up phoneme mapping and exhibits greater negativity when incoming phonemes mismatch a phonological expectation. The N400 is a negative-going central/posterior component peaking at approximately 400 ms post-stimulus onset. It reflects lexical-semantic processing and exhibits greater negativity when incoming semantic or lexical content differs from expectation (Connolly & Phillips, 1994).

ERPs during picture/spoken-word matching have revealed differences in processing cohorts vs. rhymes (Desroches, Newman & Joanisse, 2009). Cohort mismatches (e.g., see *cone* hear comb) result in no modulation of the PMN and late increased N400 negativity compared to unrelated and rhyme mismatches. These large late N400s for cohort mismatches reflect the simultaneous mismatch of phonology and lexicality, and the effortful processing required when initial bottom-up information matches an expectation. In contrast, rhyme mismatches (e.g., see *cone* hear *bone*) elicit a PMN and a reduced N400. The data suggest rhyme competitors are activated via top-down connections, leading to facilitated recognition of rhyme mismatches. The differential modulation of the PMN and N400 for cohorts and rhymes have been robust across typical development in English (Desroches et al., 2009; Desroches, Newman,

CAMBRIDGE
UNIVERSITY PRESS

Robertson & Joanisse, 2013; Malins, Desroches, Robertson, Newman, Archibald & Joanisse, 2013), and in other languages (i.e., Mandarin, rhyme and tonal mismatches showed similar facilitation; Malins & Joanisse, 2012; Malins, Gao, Tao, Booth, Shu, Joanisse, Liu & Desroches, 2014).

The dynamics of bilingual language processing

Several models have been proposed to explain language processing in bilinguals. For example, the Bilingual Language Interaction Network for Comprehension of Speech Model (BLINCS; Shook & Marian, 2013) has interconnected levels of processing that are dynamic and interactive. When activation occurs at the conceptual level, lexical options from both languages could be subsequently activated. Lexical representations are structured via self-organizing maps: that is, words within a language tend to cluster together. However, words that overlap across languages are situated closer together relative to lexical items that do not overlap. BLINCS suggests that language co-activation occurs during processing, such that within-language and between-language cohorts and rhymes compete via feedback.

There is ample evidence that bilinguals activate both languages when processing a single language (e.g., Dijkstra, Grainger & van Heuven, 1999; Duyck, 2005; Friesen, Jared & Haigh, 2014; Haigh & Jared, 2007; Jared & Kroll, 2001; Hermans, Bongaerts, De Bot & Schreuder, 1998; Marian & Spivey, 2003a; Spivey & Marian, 1999). Cross-language phonological competition has been revealed with eye-tracking using the visual world paradigm (Spivey & Marian, 1999; Marian & Spivey, 2003a, 2003b). For example, when Russian–English bilinguals were asked to move a *marker* (in their L2) in an array of objects, there was a greater proportion of looks to a *stamp* (*marka* in Russian) vs. unrelated objects. In contrast, monolinguals treated cross-language cohorts as unrelated distractors (Marian & Spivey, 2003b). This cross-language effect was robust in bilinguals' second language and of similar magnitude to within-language competitors (Marian & Spivey, 2003a). However, an impact of the second language on the first language was not observed, which the authors attributed to weaker activation of the second language due to lower proficiency in late bilinguals. Marian and Spivey's findings suggest bilinguals process between-language and within-language cohorts similarly; if activation is sufficiently strong, cross-language cohorts compete for recognition.

Very few studies have examined the neural basis of L2 lexical activation during L1 processing. Recently, Bobb and colleagues (Bobb, Von Holzen, Mayor, Mani & Carreiras, 2020) examined whether Spanish–Basque bilinguals activate labels from both languages when presented with pictures during an L1 task. In critical trials, the picture was followed by either an unrelated L1 auditory word or an L1 word that rhymed with the L2 label of the picture. ERPs were time-locked to the onset of the auditory word. Results indicate that labels for the picture in both languages were activated because bilinguals responded differently to cross-language rhymes than to unrelated word. However, the interpretation of the temporal dynamics in this dataset was challenging because, unexpectedly, the rhyme condition produced greater negativity than the unrelated condition.

Friesen and colleagues (Friesen, Chung-Fat-Yim & Bialystok, 2016; Study 2) investigated cross-language cohort effects. English–French bilinguals saw two pictures and simultaneously heard an English auditory cue; the task was to identify which picture matched the spoken word. The phonological competitor shared cross-language onset with the target. ERPs did not differ between this condition and the unrelated condition. Since stimuli were presented simultaneously, there may not have been sufficient time for cross-language phonological competition to occur. Thus, further investigation of the neural bases of cross-language competition is warranted using a paradigm that has established clear effects with monolinguals.

The current study

We used ERPs to examine spoken word recognition in English–French bilinguals and English monolinguals during a picture/spoken-word matching task (adapted from monolingual research; Desroches et al., 2009; Desroches et al., 2013; Malins et al., 2013, 2014). We included match (BEACH-*beach*), unrelated mismatch (BEACH-*tack*), and second language cohort mismatch (L2-cohort: BEACH-*plaid*; which overlaps with *plage*, the French word for *beach*) trials. Past research with monolinguals showed that within-language cohort mismatches resulted in no modulation of the PMN and larger late-going N400s, reflecting the initial matching of the bottom-up input to the phonological expectation (e.g., *cone* and *comb* do not mismatch until the final phoneme). Thus, if between-language cohorts behave like within-language cohorts on this task, this would require the bilingual to generate an explicit phonological expectation for the target picture in both of their languages (e.g., they expect to hear either *beach* or *plage*). However, phonological representations can be activated even if they are not explicitly expected, as has been observed for rhyme and tonal mismatches within-language. That is, even when a participant develops a strong target expectation, top-down connections from that target lead to activation of competitors, marked by reduced N400s to these types of mismatches. It is possible that L2-cohort mismatches might behave similarly: participants develop an explicit expectation for the L1 target phonology; however, the L2 target phonology is nevertheless activated via top-down representations. In this case we would expect that L2-cohort mismatches would result in an increased PMN compared to matches (reflecting the mismatch between the bottom-up input and phonological expectation of the L1 target), and a reduced N400 compared to unrelated mismatches, reflecting the top-down activation of the L2 phonology and its overlap with the spoken word.

Method

Participants

Data was collected from 40 right-handed participants who self-identified as functionally monolingual or as English–French bilinguals. Data was excluded from 2 bilinguals who named fewer than 70% of the stimuli in French, and 2 monolinguals who had excessive EEG artifacts. Data were also excluded from 3 participants who self-identified as monolinguals but successfully named more than 40% of the stimuli in French. The final sample included 17 bilinguals ($M_{age} = 24.1$, $SD_{age} = 3.5$; 11 females) and 16 monolinguals ($M_{age} = 24.1$, $SD_{age} = 4.0$; 10 females).

Table 1 includes self-report data on age of language acquisition, the current language use and language proficiency ratings for each language for each group. It reports the number of years students were educated solely in English and the number of years students took at least one French course. Note, monolinguals were required to take a French course from middle school

Table 1. Participants' Self-report on Language Experience and French Picture Naming Scores

	Monolinguals Mean (SD)	Bilinguals Mean (SD)
Self-Report		
English		
Age of Acquisition	1.3 (1.1)	1.3 (1.3)
Overall Current Use (%)	97.6 (5.4)	80.9 (14.0)
Proficiency Rating (out of 10)	9.2 (2.1)	9.9 (0.5)
Years of Schooling with English only	11.7 (4.1)	2.3 (1.9)
French		
Age of Acquisition	10.2 (4.4)	2.6 (2.1)
Overall Current Use (%)	1.0 (2.7)	18.8 (14.0)
Proficiency Rating (out of 10)	2.5 (1.6)	8.5 (1.3)
Years of Schooling with French classes	4.0 (2.5)	13.9 (1.6)
French Picture Naming Accuracy	14.6 (14.1)	89.5 (9.7)

Table 2. Mean and Standard Error for Accuracy and Reaction Time.

		Accuracy	Reaction Time
Bilinguals	Match	98.5% (0.5)	856.03 (24.41)
	L2-Cohort	99.2% (0.4)	945.84 (23.86)
	Unrelated	99.8% (0.2)	918.49 (26.04)
Monolinguals	Match	98.1% (0.5)	843.12 (22.07)
	L2-Cohort	99.8% (0.4)	899.22 (24.59)
	Unrelated	99.8% (0.2)	898.53 (26.83)

(age 10) to the first year of high school while at English school (4 years on average). Bilinguals only had on average 2.3 years of education without any French instruction (often at University). On average, they spent 13.9 years with instruction in French for all core subjects. Importantly, bilinguals and monolinguals did not differ on their age of English acquisition, $t(29) = -0.07$, *ns*, or their self-rated English Proficiency, $t(29) = 1.21$, *ns*. They did differ significantly on the age of French acquisition, $t(29) = 6.05$, $p < .001$, their self-rated French ability, $t(29) = 11.32$, $p < .001$, and on the French picture naming accuracy, the objective measure of French proficiency, $t(29)$ 17.88, $p < .001$.

Materials and procedure

For the picture-word matching task (E-prime-2), each picture (e.g., beach) was paired with three monosyllabic auditory stimuli: a match (e.g., *beach*), an unrelated mismatch (e.g., *tack*) and an L2-cohort competitor mismatch (L2-cohort, e.g., *plaid*). L2-cohort competitors shared cross-language onset with the correct word (i.e., *plaid* is unrelated to *beach* in English; however, it sounds like the French word for *beach*, *plage*). See the Appendix for the list of items (Supplementary Materials). Pictures were color stock images on a white background (27" monitor; resolution,1920 x 1080), and spoken words were presented over speakers (recorded by an adult female English monolingual;16 bits, 44,100 Hz, normalized).

There were 232 trials with equal proportions of matches and mismatches. There were 29 critical trials for each condition (match, L2-cohort, unrelated). Frequency did not differ across conditions (HAL, Lund & Burgess, 1996; Match: 40442.8, L2-Cohort: 27591.0, Unrelated: 27591.0, $F(2,56) = 0.5$, $p = .65$). Filler trials, which had no L2-cohort competitors and were not analyzed, were included to ensure equal proportions of matches and mismatches (87 match and 58 unrelated fillers). On each trial, a fixation-cross appeared for 250 ms, followed by the picture. After 1500 ms, a spoken word was presented while the picture remained on screen. Participants indicated whether the spoken word matched the picture with a button press ("yes," left index finger; "no," right index finger). They were allowed 2500 ms from the onset of the spoken word to respond, with 1000 ms between the response and the next trial.

When bilingual participants were recruited, they were informed that they needed to be proficient in both English and French. However, when greeted at the lab, no mention of French was made. Prior to the experiment, participants were asked to name each picture in English to ensure they would develop strong expectations of that label. If a different label was provided, corrective feedback was given (e.g., saying *flower* instead of *rose*). At no point prior to the experiment were participants explicitly cued to the French words for the items. Participants performed six practice trials to ensure that they understood the procedure. Following the experiment, we assessed participants' French knowledge by asking them to report the French name for each picture, and via a language history questionnaire.

Electrophysiological recording

Electroencephalogram (EEG) was recorded at 500 Hz (32 Ag/AgCl active electrodes; actiCHamp, Brain Products GmbH, Gilching, Germany, with a standard 10/20 placement), digitally referenced to Fz (re-referenced TP9/TP10, approx. mastoids) with impedances kept below 20 kΩ. Data (analyzed with Brain Analyzer 2.0) was filtered off-line (60 Hz notch; 0.01 Hz, 24 dB/Oct zero phase shift high pass). Data were segmented (-200 to 800 ms) and trials were baseline corrected to the pre-stimulus interval. Blinks were removed using ICA ocular correction (one component per participant on average). Trials with voltages +/- 85 µV on the sites of interest identified by past research (Fz/Cz/Pz) were removed (on channels of interest, to avoid removing trials with high voltages only on channels not of interest). Data were averaged for each participant for each condition. Participants had an average of 25.4, 26.6, and 26.3 trials included for the match, L2-cohort, and unrelated conditions, respectively. Mean amplitudes were calculated for intervals of interest, selected based on both visual inspection of the data and on past research (N400: 350–500 ms; PMN: 300–350 ms). After analyses were complete, we applied a low pass filter for visualization (30 Hz, 24 dB/Oct).

Results

This task is designed to be easy, and like past studies, accuracy was at ceiling (see Table 2). Also like past studies, there was an effect of condition ($F(2,62) = 30.50$, p < .001), with match trials showing faster latencies than mismatches ($ps < .001$, Bonferroni corrected). No significant interactions or group effects were observed for either
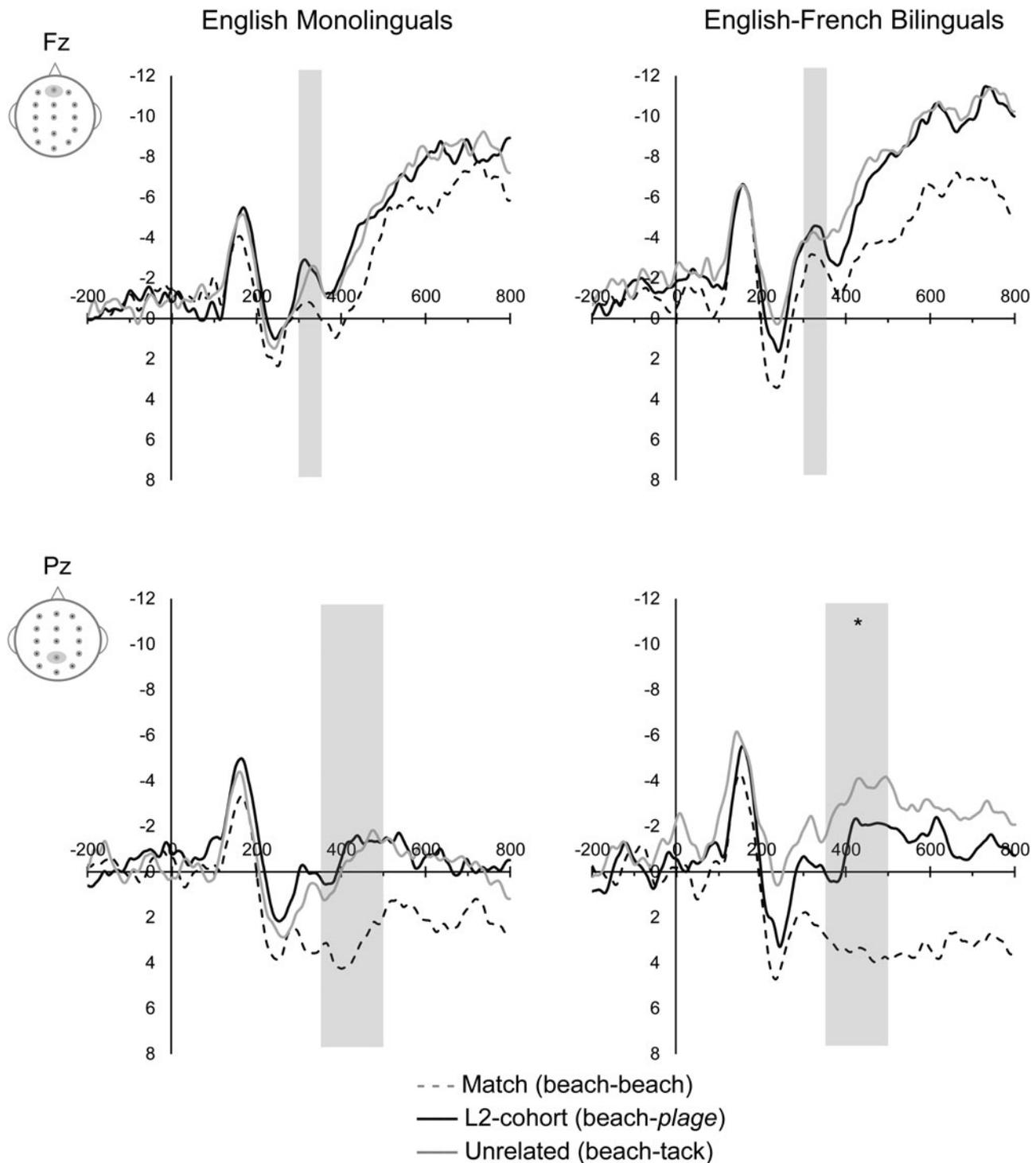
Fig. 1. Average waveforms for mismatch conditions compared to match conditions during the picture-spoken word matching task. Grey bars mark the PMN (at Fz) and N400 (at Pz) intervals. Asterisked results highlight an N400 reduction to L2-cohort mismatches for French-English bilinguals, absent in English monolinguals.

measure; thus, we focused on analysis of the electrophysiological data.

3.1 N400

The N400 was analyzed from 350–500 ms at posterior sites (Figure 1). A 3-site (Pz, P3, P4) by 3 condition (match, L2-cohort, unrelated), by 2 group (bilingual, monolingual)

repeated measures ANOVA revealed a main effect of condition ($F(2,62) = 35.56$, $p < .001$, $\eta^2 = .534$). Posthoc comparisons revealed that both the L2-cohort and the unrelated mismatch had significantly more negative amplitude than the match condition ($p < .05$, Bonferroni corrected). Overall, no differences emerged between mismatch conditions. However, this main effect was qualified by a group by condition interaction that approached significance ($F(2,62) = 3.03$, $p = .061$, $\eta^2 = .089$). We had strong a

priori predictions related to group differences between the mismatch conditions, which may have been masked in this ANOVA due to the inclusion of the match condition. To pursue our a priori predictions regarding group effects on the unrelated and L2-cohort mismatches, a 3-site (Pz, P3, P4) by 2 condition (L2-cohort, unrelated), by 2 group (bilingual, monolingual) repeated measures ANOVA was performed. This ANOVA revealed a significant interaction between group and condition ($F(1,31)$ = 5.39, $p < .027$, $\eta^2 = .148$, estimated power = 0.991 (G*Power 3.1, Faul, Erdfelder, Buchner & Lang, 2009). Bonferroni pairwise comparisons within each group revealed that for bilinguals, N400s were significantly reduced for L2-cohort vs. unrelated mismatches ($p = .013$). No differences were observed between conditions for monolinguals ($p = .507$).

PMN

The PMN was analyzed from 300–350 ms at F3, Fz, F4; C3, Cz, and C4 (Figure 1). These sites were selected since the PMN typically has a frontal-central distribution (Lewendon et al., 2020). A 2-row (frontal, central) by 3 column (left, center, right) by 3 condition (match, L2-cohort, unrelated), by 2 group (bilingual, monolingual) repeated measures ANOVA revealed a main effect of condition ($F(2,62) = 8.11$, $p = .001$, $\eta^2 = .207$). Bonferroni pairwise comparisons revealed significant PMNs, marked by greater negativity for each of the mismatch conditions compared to match ($ps < .005$). No difference in the PMN was observed between L2-cohort vs. unrelated mismatches ($p = 1.0$). There were no significant interactions that included both group and condition (all $Fs < 1.1$).

Discussion

We used a picture/spoken-word matching paradigm to examine cross-language cohort competition in English–French bilinguals vs. monolinguals. We found group differences in N400 modulations to L2-cohorts vs. unrelated mismatches. Specifically, bilinguals showed reduced N400s to L2-cohorts vs. unrelated mismatches, while monolinguals treated both conditions similarly. Importantly, the results revealed that bilinguals activated the pictures' labels in both their languages when performing a task in L1. Moreover, they help to specify the nature of cross-language connections, indicating that top-down cross-language connections exist between lexical representations and phonology.

Our study revealed that between-language cohort mismatches resulted in large PMNs and reductions in the N400. These results contrast the reduced PMNs and the large late-going N400s observed in previous studies of within-language cohorts (Desroches et al., 2009; Malins et al., 2013). Thus, although our results indicate between-language cohort activation occurred, the nature of this cohort activation differs from within-language cohort effects. These results may be accounted for by considering an expectancy generation explanation. In a within-language task, listeners generate a phonological expectation from the picture in the target language (e.g., *beach*). When a within-language cohort is presented (e.g., *bead*), the bottom-up input is initially consistent with the expectation resulting in the reduced PMN. At the point of uniqueness, the listener must engage in effortful processing to identify the cohort as a mismatch, resulting in greater N400 negativity. For between-language cohorts, our results indicate that bilinguals do not develop an explicit phonological expectation of the French word (e.g., *plage*), thus the between-language cohort

(e.g., *plaid*) does not elicit the same pattern of results as a within-language competitor on this task.

Nevertheless, the results indicate that bilinguals do activate L2 phonological representations for targets on this task, marked by reduced N400s to L2-mismatches. This pattern of response for L2-cohorts for bilinguals is strikingly similar to what has been observed for rhymes and tones, which also show reduced N400 effects (e.g., Desroches et al., 2009; Malins et al., 2013). Desroches and colleagues suggested that this N400 reduction occurs because top-down connections from the expectation to phonologically similar items leads to facilitated recognition of rhyme mismatches. We suggest that something similar is going on for the L2-competitors for bilinguals on this task: seeing a picture of a *beach* leads to top-down activation of the phonemes in *plage,* even though *plage* is not expected. For L2-mismatch trials, hearing *plaid* after seeing *beach* leads to reduced N400 responses compared to unrelated, which could only happen if L2 phonology was activated.

Despite the differences between within-language and between-language cohort effects, we do not consider our findings to be inconsistent with those of Marian and colleagues (Spivey & Marian, 1999; Marian & Spivey, 2003a, 2003b), but rather complementary. For cohorts in the visual world paradigm there is explicit competition: listeners select between objects in a display, and both target and cohort pictures match the initial bottom-up input. In contrast, our design does not require the listener to select between objects. Instead, our paradigm investigates how top-down expectations may influence and interact with subsequent bottom-up processing. Importantly, this effect, marked by reduced N400s compared to unrelated mismatches, has been observed using rhyme mismatches in English, rhyme and tonal mismatches in Mandarin, and now using cross-language cohorts in English–French bilinguals.

Our results provide strong evidence of language co-activation consistent with models of written and spoken bilingual word recognition (e.g., Dijkstra, Wahl, Buytenhuijs, van Halem, Al-Jibouri, de Korte & Rekké, 2019; Shook & Marian, 2013). However, current bilingual models underspecify the role of language expectations in word recognition. BLINCS (Shook & Marian, 2013) is driven by bottom-up relationships between words and does not specify a global language system to account for language expectancies. Similarly, the Multilink model (Dijkstra et al., 2019) does not specify a top-down role for language nodes to generate expectations about single or dual language input, but rather explains effects at the lexical level through resting activation. This distinction is important, given our results suggest that language expectations themselves impact the nature of cross-lexical activation in our paradigm. The current study demonstrates the importance of expectancy in how language co-activation is manifested at the neural level and this will need to be accounted for in future models of bilingual activation.

References

Allopenna PD, Magnuson JS and Tanenhaus MK (1998) Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *Journal of Memory and Language* **38**, 419–439. doi:10.1006/jmla.1997.2558

Bobb SC, Von Holzen K, Mayor J, Mani N and Carreiras M (2020) Co-activation of the L2 during L1 auditory processing: An ERP cross-modal priming study. *Brain and Language* **203**, 13. doi:http://dx.doi.org.proxy1.lib.uwo.ca/10.1016/j.bandl.2019.104739

Connolly JF and Phillips NA (1994) Event-related potential components reflect phonological and semantic processing of the terminal word of spoken sentences. *Journal of Cognitive Neuroscience* **6**, 256–266. doi:10.1162/jocn.1994.6.3.256

Desroches AS, Joanisse MF and Robertson EK (2006) Specific phonological impairments in dyslexia revealed by eyetracking. *Cognition* **100**, B32-B42. doi:10.1016/j.cognition.2005.09.001

Desroches AS, Newman RL and Joanisse MF (2009) Investigating the time course of spoken word recognition: electrophysiological evidence for the influences of phonological similarity. *Journal of Cognitive Neuroscience* **21**, 1893–906. doi:10.1162/jocn.2008.21142

Desroches AS, Newman R.L, Robertson E.K and Joanisse M.F (2013) Electrophysiological indices of phonological processing impairments in dyslexia. *Journal of Speech Language and Hearing Research* **56**, 1–15. doi: 10.1044/1092-4388(2012/10-0351)

Dijkstra T, Grainger J and van Heuven WJB (1999) Recognition of cognates and interlingual homographs: The neglected role of phonology. *Journal of Memory and Language* **41**, 496–518. doi: 10.1006/jmla.1999.2654

Dijkstra T, Wahl A, Buytenhuijs F, van Halem N, Al-Jibouri Z, de Korte M and Rekké S (2019) Multilink: A computational model for bilingual word recognition and word translation. *Bilingualism: Language and Cognition* **22**, 657–679. doi: 10.1017/S1366728918000287

Duyck W (2005) Translation and associative priming with cross-lingual pseudohomophones: Evidence for nonselective phonological activation in bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition* **31**, 1340–1359. doi:10.1037/0278-7393.31.6.1340

Faul F, Erdfelder E, Buchner A and Lang A.-G (2009) Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods* **41**, 1149–1160. doi:10.3758/BRM.41.4.1149

Friesen DC, Chung-Fat-Yim A and Bialystok E (2016) Lexical selection differences between monolingual and bilingual listeners. *Brain and Language* **152**, 1–13. doi:10.1016/j.bandl.2015.11.001

Friesen DC, Jared D and Haigh CA (2014) Phonological processing dynamics in bilingual word naming. *Canadian Journal of Experimental Psychology* **68**, 179–193. doi: 10.1037/cep0000026

Green DW (1998) Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition* **1**, 67–81. doi: https://doi.org/10.1017/S1366728998000133

Haigh CA and Jared D (2007) The activation of phonological representations by bilinguals while reading silently: Evidence from interlingual homophones. *Journal of Experimental Psychology: Learning, Memory, & Cognition* **33**, 623–644. doi: 10.1037/0278-7393.33.4.623

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? *Bilingualism: Language and Cognition* **1**, 213–229. doi:10.1017/S1366728998000364

Jared D and Kroll J (2001) Do bilinguals activate phonological representations in one or both of their languages when naming words? *Journal of Memory and Language* **44**, 2–31. doi: 10.1006/jmla.2000.2747

Lewendon J, Mortimore L and Egan C (2020) The Phonological Mapping (Mismatch) Negativity: History, Inconsistency, and Future Direction. *Frontiers in psychology* **11**, 1967. https://doi.org/10.3389/fpsyg.2020.01967

Lund K and Burgess C (1996) Producing high-dimensional semantic spaces from lexical co-occurrence. *Behavior Research Methods, Instruments, & Computers* **28**, 203–208. https://doi.org/10.3758/BF03204766

Malins JG, Desroches AS, Robertson EK, Newman RL, Archibald LMD and Joanisse MF (2013) ERPs reveal the temporal dynamics of auditory word recognition in specific language impairment. *Developmental Cognitive Neuroscience* **5**, 134–148. doi:10.1016/j.dcn.2013.02.005

Malins JG, Gao D, Tao R, Booth JR, Shu H, Joanisse MF, Liu L and Desroches AS (2014) Developmental differences in the influence of phonological similarity on spoken word processing in mandarin Chinese. *Brain and Language* **138**, 38–50. doi:10.1016/j.bandl.2014.09.002

Malins JG and Joanisse MF (2012) Setting the tone: An ERP investigation of the influences of phonological similarity on spoken word recognition in mandarin Chinese. *Neuropsychologia* **50**, 2032–2043. doi:10.1016/j.neuropsychologia.2012.05.002

Marian V and Spivey M (2003a) Competing activation in bilingual language processing: Within- and between-language competition. *Bilingualism: Language and Cognition* **6**, 97–115. doi:10.1017/S1366728903001068

Marian V and Spivey M (2003b) Bilingual and monolingual processing of competing lexical items. *Applied Psycholinguistics* **24**, 173–193. doi:10.1017/S0142716403000092

McClelland JL and Elman JL (1986) The TRACE model of speech perception. *Cognitive Psychology* **18**, 1–86. doi:10.1016/0010-0285(86)90015-0

Shook A and Marian V (2013) The bilingual language interaction network for comprehension of speech. *Bilingualism: Language and Cognition* **16**, 304–324. doi:10.1017/S1366728912000466

Spivey MJ and Marian V (1999) Cross talk between native and second languages: Partial activation of an irrelevant lexicon. *Psychological Science* **10**, 281–284. doi:10.1111/1467-9280.00151