Lexical recognition processes in L2-dominant bilingualism

Laurence Bruggeman[1]* and Anne Cutler[1,2]†

[1]The MARCS Institute for Brain, Behaviour and Development and The ARC Centre of Excellence for the Dynamics of Language, Western Sydney University, Penrith, NSW, Australia, [2]Max Planck Institute for Psycholinguistics, Nijmegen, Netherlands

To comprehend speech, listeners must resolve competition between potential candidate words. In second-language (L2) listening such competition may be inflated by spurious activation; the onsets of "reggae" and "legacy" may both activate "leg" for Japanese listeners, or the rhymes of "adapt" and "adept" may activate "apt" for Dutch listeners, while only one in each pair triggers competition for L1 listeners. Using eyetracking with L2-dominant bilingual emigrants, we directly compared within-language L1 and L2 lexical activation and competition in the same individuals. For these listeners, activation patterns did not differ across languages. Unexpectedly, however, we observed onset competition in both languages but rhyme competition in the L2 only (although the same stimuli elicited rhyme competition for control listeners in both languages). This suggests that L1 rhyme competition may disappear after long-time immersion in an L2 environment.

KEYWORDS

spoken-word recognition, lexical competition, bilingualism, dominance, emigrants, L2 listening

1 Introduction

Spoken language typically does not provide clear boundaries between words. This means that, to understand what is being said, listeners must segment the continuous stream of speech into the separate words it contains. This segmentation process does not launch only once some criterion of sufficiency has been attained; it starts as soon as the first speech sounds become available. As the speech signal unfolds, and more information about what is being said becomes available, words that fully or partially correspond to the speech input are activated and compete for recognition, to be then either ruled out as viable candidates, or recognised once they have "won" the lexical competition (for a review, see Eisner and McQueen, 2018).

Unsurprisingly, words that overlap at the beginning (so-called onset competitors, such as *button* and *butter*) become simultaneously active if their overlap portion is heard, and effectively compete with one another for that portion of the input. This onset competition is predicted by all models of spoken-word recognition (e.g., Cohort: Marslen-Wilson and Welsh, 1978; TRACE: McClelland and Elman, 1986; Shortlist: Norris, 1994; Shortlist B: Norris and McQueen, 2008) and has been repeatedly attested in experiments (e.g., Allopenna et al., 1998; McQueen and Viebahn, 2007). However, the ends of words also affect lexical activation (e.g., Taft and Hambly, 1986; Slowiaczek et al., 1987; Goodman and Huttenlocher, 1988; Connine et al., 1993), and words that only overlap with the <u>end</u> of a spoken word (known as rhyme competitors, e.g., *bumper* and *jumper*) have been shown to compete for recognition as well, whereby this competition occurs later and is generally less strong than that of onset competitors (e.g., Allopenna et al., 1998; McQueen and Viebahn, 2007) but may increase somewhat in strength under conditions of potential noise interference (McQueen and Huettig, 2012; Brouwer and Bradlow, 2016).

Second language (L2) listeners face many difficulties during spoken-word recognition that native (L1) listeners do not have to contend with. This makes L2 listening harder than L1 listening (e.g., Clopper and Bradlow, 2009; Garcia Lecumberri et al., 2010; Bregman and Creel, 2014), and at least some of these difficulties concern the processes of lexical activation and competition. Bilinguals' misperception of L2 phonemes, for instance, may cause activation of additional (incorrect) lexical candidates. Thus difficulty distinguishing between English /l/ and /r/ means that for many Japanese listeners, hearing the English fragment *rock*-temporarily activates *locker* as well as *rocket* (Cutler et al., 2006). Likewise, Dutch L2 listeners of English – many of whom struggle with the distinction between English phonemes /æ/ and /ɛ/ – experience activation of *deficit* as well as *daffodil* upon hearing *daf*-, or *pencil* as well as *panda* upon hearing *pan*- (Weber and Cutler, 2004; Broersma, 2012).

This prolonged activation can eventually be resolved, and spurious lexical candidates deactivated, once further phonemes are recognised (*pand*- no longer matches the onset of *pencil*, leading to deactivation of *pencil*). However, false competitors can also be exceptionally hard to get rid of, because stored lexical representations are constructed not only from speech experience; they also reflect reading experience (i.e., orthography) and information from teachers, and thus include knowledge of distinctions that actually are not perceived in the speech signal (e.g., Escudero et al., 2008). When Dutch listeners hear the English word *daffodil*, they tend to misperceive the first vowel /æ/ as /ɛ/. This leads to the activation of the false competitor *deaf*. Since the listeners' lexical representation of *daffodil* contains a vowel which is marked as <u>not</u> the same as the vowel of *deaf*, the word *daffodil* does <u>not</u> get activated. Because *deaf* is what the listener hearing *daffodil* first perceives, *deaf* remains active even after all segments of *daffodil* have been heard, because *daffodil*, due to that marking, simply does not compete with it (Broersma and Cutler, 2011).

Such phonetic confusion also affects minimal pairs, so that *rice* and *lice* may sound like homophones to Japanese listeners, and *cattle* and *kettle* may become homophones for Dutch listeners (Cutler and Otake, 2004; Broersma, 2012; Díaz et al., 2012). In these cases, with deactivation of the spurious lexical candidates not possible, L2 listeners are left to cope with more lexical competitors than L1 listeners would experience given the same spoken input. As if this were not enough, additional competition may also arise when misperception of L2 phonemes causes L2 listeners to misinterpret nonsense words as real words (Broersma and Cutler, 2011).

On top of this, co-activation of lexical items from the *L1* also increases the number of activated and competing lexical candidates in L2 listening. Dutch L2 listeners (but not English L1 listeners) experienced significant competition from Dutch *deksel* ("lid") when hearing English *desk* (Weber and Cutler, 2004). Similarly, upon hearing the English word *marker*, Russian students at an American university experienced competition not only from English *marbles*, but also from Russian *marka* ("stamp"; Marian and Spivey, 2003).

Bilinguals thus have to contend with a larger set of lexical competitors in the L2, from both their L2 and their L1. For some bilinguals—in particular those with high L2 proficiency living in an environment where the L2 is the language of daily use—the reverse is also the case: lexical candidates from the L2 are co-activated during L1 listening, leading to a larger competitor set in L1 listening as well (e.g., Spivey and Marian, 1999; Blumenfeld and Marian, 2007, 2013; Lagrou et al., 2011). As an increase in the number of competitors slows word recognition (Luce et al., 1990; Norris et al., 1995), the larger competitor sets might be expected to slow bilingual listeners down. As yet, however, very little is known about how L1 and L2 lexical processing compare in one and the same listener, although what evidence there is suggests that the neurophysiological patterns of an individual's processing of L1 vs. L2 run a similar course even when the two languages are phonologically quite different (Xue et al., 2020).

In the study reported here, we address the issues of within-language lexical activation and competition in L1 and L2 listening in bilingual listeners. We use a language pair with high phonological similarity (Dutch, English) and ask whether these lexical processes occur in the same way for each language, or whether under conditions of phonological similarity one or the other language may have an advantage. In particular, we further ask whether the patterns of activation and competition for onset vs. rhyme competitors differ across L1 and L2.

Our participant population consists of Dutch-English bilingual emigrants living in Australia, an L2 immersion environment. This is atypical for studies of bilingual speech perception where L2 is usually the language spoken less often and with lesser proficiency. For this population, L2 is the more frequently used language and in most situations the dominant language, so that we are able to ask further whether listeners who are effectively L2-dominant still show the typical lexical competition processes in their L1. In equivalent experiments in the bilingual emigrants' L1 and L2, we compare the lexical activation and competition processes in each language by using the visual world paradigm (Tanenhaus et al., 1995; Allopenna et al., 1998; McQueen and Viebahn, 2007; for a review, see Huettig et al., 2011). This method has proven very successful in capturing the time course of spoken word recognition and has been widely used to investigate lexical competition processes, including in L2 listeners (e.g., Ju and Luce, 2004; Weber and Cutler, 2004; Blumenfeld and Marian, 2007; Chambers and Cooke, 2009; Canseco-Gonzalez et al., 2010; Lagrou et al., 2013).

2 Experiment I

2.1 Participants

Twenty members of the Dutch emigrant community in Sydney, Australia (aged 27-73 years, $M = 51.6$, $SD = 15.2$; 12 females, 8 males), were paid to take part in two experimental sessions. They were recruited on social media and via advertisements at Dutch venues in Sydney (e.g., consulate-general of the Netherlands, Dutch supermarket). Data from three further emigrants were excluded due to eyetracker calibration difficulties in both sessions (2), or failure to complete the experiment (1). All included participants were born and raised in the Netherlands, were native speakers of Dutch and had migrated to Australia as adults (mean age at migration = 29.1 years, $SD = 8.4$, range: 18-52). Their mean length of residence in Australia was 22.3 years ($SD = 16.2$). Although nowadays all children in the Netherlands are taught English in

school from around age 10, this has not always been the case. Hence, several participants in the present study started learning English from a later age, and one did not know any English at all before emigrating. On average, participants received 6.2 years (range: 0–10 years) of formal English instruction, starting at a mean age of 11.4 years (range: 10–16 years). Participants' proficiency in L1 and L2 was assessed with the Dutch and English versions of the LexTALE (Lemhöfer and Broersma, 2012). This brief vocabulary-based test showed that the participants were highly proficient in their L2, English, with a mean score of 94.4% (SD = 5.2, range: 84-100). Despite migration to a predominantly English-speaking country, participants had also maintained high proficiency in their native Dutch, as evidenced by a mean LexTALE score of 92.4% (SD = 5.5, range: 80-100). We also asked participants—as part of a background questionnaire they completed before the experiment—to indicate to what extent they used Dutch and English in a range of situations. This showed that all participants used the L2, English, more frequently than the L1, Dutch. The full list of situations and a tally of participants' answers can be found in Appendix A in the Supplementary material. All participants had normal or corrected-to-normal vision. Each participant provided written informed consent before the start of the experiment.

2.2 Design and materials

Two versions of the experiment were created. In the L1 version, all materials were in Dutch; in the L2 version, they were in English. A female native speaker of Dutch (for the L1 version) or Australian English (for the L2 version) recorded three sets of 40 sentences (one set each for the Onset and Rhyme conditions, plus a filler set; see Appendices B, C in the Supplementary material) in a sound-attenuated booth, using Adobe Audition and a sampling rate of 44.1 kHz. Each sentence contained a critical word (e.g., *penalty*) that occurred in mid-sentence and was not easily predictable from the preceding context (e.g., *She was not sure whether the penalty for the other team was totally fair*). Sentences were read aloud with neutral intonation, and the speaker was unaware of the presence or identity of the critical words. To achieve equivalent loudness across the stimuli, the amplitude of each sentence was root-mean-square (RMS) standardised in Praat (Boersma and Weenink, 2013). Each sentence was paired with a visual display containing four black-and-white line drawings (see Supplementary Figures D.1 and D.2 for an example display in each language). Three of the four drawings were distractors that were phonologically and semantically unrelated to the critical word. The fourth drawing either matched to the critical word itself (filler trials), or to a competitor that overlapped with the critical word in its onset (Onset condition; e.g., *pencil* as competitor for *penalty*) or in its rhyme (Rhyme condition; e.g., *bikini* as competitor for *zucchini*). A list of critical words and their competitors is shown in Appendix E in the Supplementary material. Thus, while the filler condition contained a "target" drawing depicting the critical word, the competitor conditions did not (making for a so-called "target absent" design; Huettig and Altmann, 2005; McQueen and Huettig, 2012). In all conditions in both language versions, competitor and

distractor pictures were counterbalanced across the four positions on the screen.

The critical words in the Onset condition comprised on average 2.5 syllables in the L1 version of the experiment, and 2.1 syllables in the L2 version. The mean number of syllables for critical words in the Rhyme condition was 1.25 in the L1, and 1.5 in the L2 version. The mean overlap of onset competitors with the critical word in the spoken sentence was 3.8 phonemes for the L1, and 3.5 phonemes for the L2 version. Mean word frequency of these competitors based on the CELEX lexical database (Baayen et al., 1995) was 10.3 and 16.7 per million words, in the L1 and L2 materials respectively. Rhyme competitors had a mean overlap of 3.2 phonemes in the L1, and 3.5 phonemes in the L2 materials, and a mean CELEX word frequency of 51.8 and 20.2 per million words, respectively. As definite articles and adjectives are marked for gender in Dutch, we forestalled gender-based biases by ensuring that all referents in each display of the L1 version were grammatically compatible with articles and adjectives preceding the critical word in the spoken sentence. In the L2 version, to prevent listeners from making predictions based on the form of indefinite articles (*a/an*), which in English depends on the onset of the word that follows them, visual displays for sentences in which the critical word was preceded by an indefinite article only contained pictures for which the referent matched the used article. Further, since cross-language co-activation of lexical candidates may affect the proportion of looks that referent pictures attract from bilingual listeners, care was taken during stimulus selection to control for potential co-activation in Dutch and English: The distractors in each visual display were always phonologically unrelated to the competitor or target they occurred with, not only in the language of the experiment version, but also in the other language. For example, the competitor picture *pencil* (see Supplementary Figure D.2) was matched with distractor images of *stove, hammock*, and *gloves*. The Dutch referents for these pictures are *fornuis, hangmat*, and *handschoenen*, all of which are phonologically unlike the competitor *potlood*. In addition, as cross-language co-activation particularly affects cognates (Blumenfeld and Marian, 2007), competitors that were cognates in Dutch and English were always combined with three distractors that were cognates in both languages as well. This was the case for 17 Onset condition trials and 13 Rhyme condition trials in the L1 version of the experiment, and for 18 Onset and 14 Rhyme trials in the L2 version of the experiment.

2.3 Procedure

All participants completed both the L1 and the L2 experiment session, approximately three weeks apart (M = 18.3 days, SD = 4.7, range 12-29 days), with session order counterbalanced across participants. During the L1 session, all stimulus materials, tasks and instructions were presented in the participants' L1, Dutch; during the L2 session these were in their L2, English. All sessions were conducted by the same experimenter (the first author), who is a native speaker of Dutch and is fluent in English. The language that was spoken during each session depended on the participant's preference; for most participants, this was Dutch. The same experimental procedures were followed for both sessions.

Participants were tested individually in a sound-attenuated booth. Their eye movements were recorded using an SR Research Eyelink 1000 Tower Mount system, which held their head in a fixed position by means of a chin and forehead rest, and was placed at a viewing distance of 65 cm from a computer screen. Eye movements were sampled at a frequency of 1,000 Hz (monocular). Auditory stimuli were presented over Beyerdynamic DT770-pro headphones at a loud but comfortable level, kept constant for all participants. Written instructions were presented on the screen. They were subsequently repeated orally by the experimenter and clarified if needed. Participants were given no explicit task, other than to listen to the sentences and to keep their eyes on the screen. Before the start of the experiment, the eyetracker was calibrated and validated using a 9-point calibration grid. After every five trials, an automatic drift check was carried out and, if required, calibration was repeated. At the start of each trial, a fixation cross was displayed in the centre of the screen. Participants were instructed to look at this cross until it disappeared. Once the fixation cross had disappeared, the visual display was shown for 1 s, after which the sentence started playing while the display remained on the screen.

As the experiment reported here formed part one of a larger eyetracking study with several listening conditions, participants were presented with only 60 of the 120 available sentence-display pairs within each session (20 Onset, 20 Rhyme and 20 filler trials). Sentences were counterbalanced across participants such that half of all participants were presented with one set of 60 pairs, whereas the other half was presented with the remaining pairs. Each sentence-display pair was presented just once to each participant, and a different randomisation of the stimulus list was used for each participant. Following the eyetracking task, participants completed the LexTALE proficiency test (Lemhöfer and Broersma, 2012) in the language of the session.

2.4 Results and discussion

Data for one participant in the L1 session, and from two other participants in the L2 session were excluded from analysis due to calibration difficulties. The analyses reported here thus include data from 19 listeners in the L1 session, and 18 listeners in the L2 session.[1] All pre-processing and all analyses reported below were carried out in R (R Core Team, 2019), with the exception of the curve fitting (see below), which was done using Matlab (R2019b; The MathWorks, Inc.). To ensure integrity of the data, trials with a track loss percentage >30% were excluded from analysis. This was the case for 18 (of a total of 1,140) trials for the L1 experiment and 16 (of 1,080) trials for the L2 experiment. Mean fixation proportions to competitors and distractors (Onset and Rhyme condition) and to target and distractors (Filler trials) are shown for both experiment versions in Figure 1, plotted from the onset of the critical word. Distractor proportions were averaged over three distractors.

2.4.1 Time-course analysis of lexical activation

To compare the emigrants' looking patterns during L1 and L2 listening, non-linear curves were fitted to the competitor fixation proportions in both competitor conditions and languages, and the parameters defining these curves were analysed statistically (cf. Farris-Trimble and McMurray, 2013). This analysis method works particularly well with data of the type collected in this study, as competitor fixation proportions typically follow an asymmetric double-Gaussian curve (i.e., fixation proportions rise from one asymptote to a peak and then fall to a second asymptote). This method also enables us to draw statistical conclusions about *when* effects occur, in contrast to other methods, such as cluster-based permutation tests which do not allow any inferences about effect latency (Sassenhagen and Draschkow, 2019). As each participant was presented with only 20 trials per competitor condition in each language, individual participants' data were relatively noisy. To improve curve fitting accuracy, a jackknifing procedure (Miller et al., 1998; Ulrich and Miller, 2001) was therefore applied to the data prior to curve fitting (cf., Apfelbaum et al., 2011; Holt et al., 2021). Jackknifing involves averaging a data set across all participants except one, and repeating this, excluding a different participant each time, until each participant has been excluded once. Thus, the first subsample for each competitor condition in each language comprised the averaged fixation proportions of all participants except participant 1, the second subsample comprised the averaged proportions of all participants except participant 2, etc.

Since the time-course of competitor fixations typically follows a Gaussian-like shape, after jackknifing, six-parameter asymmetric Gaussian curves were fit to the competitor fixation proportion subsamples in Matlab (R2019b; The MathWorks, Inc.) using the curvefitter package (McMurray, 2017). Curve fits were good (mean $R = 0.96$; range: 0.79–1). Of the parameters defining asymmetric Gaussian curves, peak latency (the time point at which competitor fixations reach their peak), and peak height (the fixation proportion at the peak) best represent the strength and the time course of lexical competition. The values for these parameters were based on curve fitting to jackknifed data, so we first retrieved estimates of the actual peak latency and peak height for each individual participant (see Table 1 for mean estimates) using the method described by Smulders (2010). We subsequently fitted linear regression models to these estimates using the packages lme4 (Bates et al., 2015) and lmerTest (Kuznetsova et al., 2017) in R (R Core Team, 2019).[2] Each model included the deviation-coded fixed categorical predictors Language (L1 coded as −0.5, L2 as 0.5) and Competitor condition (Onset coded as −0.5, Rhyme as 0.5). Random intercepts and slopes for participants were added to each model as well. The results of the model fitted to peak latency as well as that fitted to peak height showed a significant effect of Competitor condition (see Tables 2, 3), indicating that fixations to onset competitors peaked earlier and higher than those to rhyme competitors. There were no significant effects of or interactions with Language, indicating no significant difference between L1 and L2 listening.

---

1   This comprises 20 unique participants in total: seventeen listeners who contributed data to both language sessions, two listeners whose data was included for the L1 session only, and one other listener whose data was included for the L2 session only.

---

2   Linear regression models fitted to the remaining curve parameters (i.e., onset and offset amplitude, and onset and offset slope) did not show any significant effects or interactions.
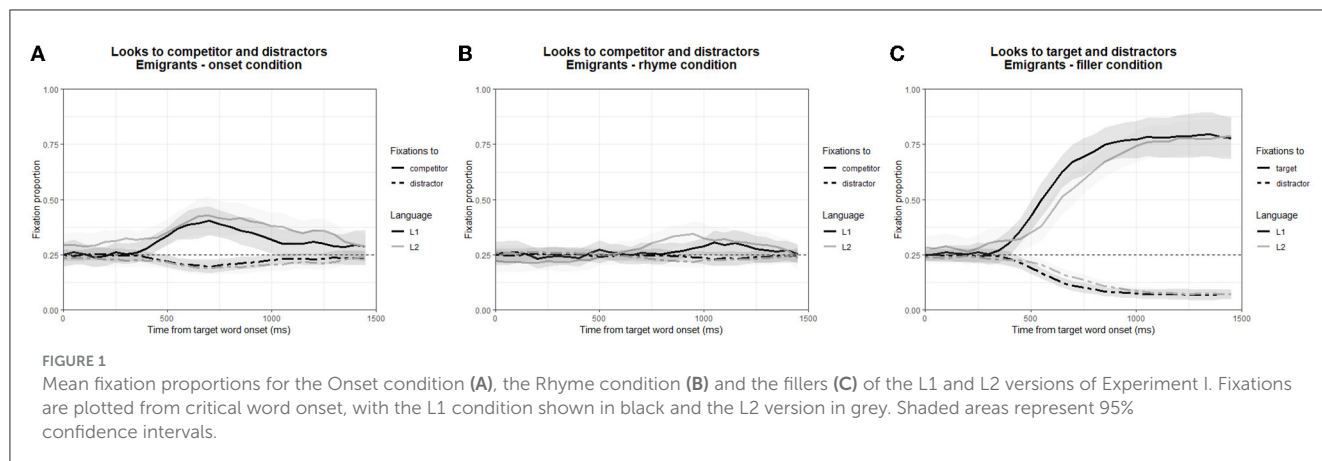
FIGURE 1
Mean fixation proportions for the Onset condition **(A)**, the Rhyme condition **(B)** and the fillers **(C)** of the L1 and L2 versions of Experiment I. Fixations are plotted from critical word onset, with the L1 condition shown in black and the L2 version in grey. Shaded areas represent 95% confidence intervals.

TABLE 1 Mean retrieved estimates for peak latency and peak height by Competitor condition and Language for participants in Experiment I.

Language	Onset condition		Rhyme condition	
	Peak latency (ms)	Peak height (fixation proportion)	Peak latency (ms)	Peak height (fixation proportion)
L1	665	0.39	1,134	0.30
L2	713	0.42	921	0.33

TABLE 2 Results of the model fitted to peak latency in Experiment I.

	β	SE	t	p
Intercept	858.83	39.39	21.804	<0.001
Language	−82.34	78.78	−1.045	0.300
Condition	337.88	78.78	4.289	<0.001
Language * condition	−261.06	157.56	−1.657	0.102

TABLE 3 Results of the model fitted to peak height in Experiment I.

	β	SE	t	p
Intercept	0.359	0.015	23.234	<0.001
Language	0.029	0.029	1.006	0.319
Condition	−0.090	0.029	−3.091	<0.001
Language * condition	0.016	0.058	0.277	0.783

2.4.2 Lexical competition

While time-course analyses provide important insights into the emigrants' lexical *activation* processes (i.e., the start of word form processing), they do not directly address lexical *competition* (i.e., the relative activation of competitor word forms compared to distractor word forms) as such. In the paradigm used here, an activated lexical candidate is considered as competing for recognition only if its referent picture attracts more looks than chance. Since each visual display contained four images, a competitor fixation proportion higher than 0.25 indicates that over a quarter of all looks to a picture were directed to the competitor picture and that the referent for the competitor picture therefore competed for recognition. Thus, to determine whether competition took place in the Onset and Rhyme conditions of the L1 and L2 experiments, we compared competitor fixation proportions in each of these conditions and languages to chance using cluster-based permutation analyses (CPA), a non-parametric statistical technique that can control for the auto-correlation that occurs in eyetracking data (Maris and Oostenveld, 2007). Although CPA does not afford any statistical conclusions about the duration of an effect, nor about its onset or offset (Sassenhagen and Draschkow, 2019), it works well for establishing the presence of an effect *somewhere* in an analysis window.

First, the appropriate analysis window was selected. Since it takes around 200 ms to initiate an eye movement (Matin et al., 1993; Luna et al., 2008), looks earlier than 200 ms after critical word onset cannot reflect processing of that word. Looking at the emigrants' looking patterns during the filler trials—which were the only trials with visual displays containing an image of the

critical word itself instead of that of a competitor word—it becomes clear that target and distractor fixation proportions do not start to diverge until around 400 ms after critical word onset (see Figure 1). This indicates that this is the earliest moment at which lexical activation takes place, so that no lexical competition is expected before this point in time. We therefore selected a one-second analysis window starting at 400 ms. We then binned the fixation data in this window in 10-ms time bins. For each time bin, we computed the test statistic, a one-sample two-tailed t-test against chance (i.e., 0.25). Since adjacent time bins are not independent of one another but are assumed to be linked to the same underlying cognitive process, we then created clusters of adjacent time bins with a significant test statistic, and computed the mass statistic for each cluster (here: the sum of all individual t-values in the cluster). Subsequently, we completed the same procedure for 1,000 mock data sets in which the fixation proportion for the competitor and for chance level were shuffled randomly for each time bin, thus creating a distribution of the cluster mass statistic under the null hypothesis (i.e., the distribution of cluster mass statistics one would expect to find if competitor fixation proportions did not differ from chance level). Finally, we compared the mass statistic for each cluster in our original data set to the null hypothesis distribution to determine their statistical significance.

The CPA showed that onset competitors were fixated significantly more than chance, in the L1 (significant cluster from 450 to 1,030 ms; cluster summed t-value: 215.93; $p <0.001$) as well as in the L2 (significant cluster from 470 to 1,380 ms; cluster summed t-value: 378.56; $p < 0.001$). In the Rhyme condition, competitors were fixated at an above-chance level in the L2 (significant cluster from 760 to 1,030 ms; cluster summed t-value: 80.54; $p < 0.001$), but, somewhat surprisingly, no significant rhyme competition was found in the L1 (all $p > 0.9$). This suggests that the bilingual emigrants experienced onset competition in both of their languages, but rhyme competition only in the L2.

While the presence of onset competition in both languages is in line with previous findings (e.g., L1: Allopenna et al., 1998; McQueen and Huettig, 2012; Brouwer and Bradlow, 2016; L2: Mercier et al., 2014; Sarrett et al., 2021), the lack of rhyme competition in the L1 is surprising, as it contrasts with the rhyme-competitor effects that have repeatedly been reported for L1 listeners (including by all the above-cited L1 studies). Before interpreting this finding, however, we must examine two factors that may, at first glance, be assumed to have played a role in these contrastive findings. The first of these relates to the age of our participants. The majority of previous evidence of rhyme competition stems from studies with undergraduate students, whereas the emigrants who participated in Experiment I were older (the youngest was 27, the oldest 73). However, onset and rhyme competition are known to pattern similarly in older and younger listeners (Ben-David et al., 2011), and the group of emigrants did experience rhyme competition in the L2, making it unlikely that the lack of rhyme competition found in the L1 was influenced by the age of some of our emigrant participants. The second factor concerns the stimulus materials, which may have prevented rhyme competition from occurring in the L1. To ensure that the experimental stimuli used in this study had the desired validity, we conducted two control experiments. The first of these, Experiment

II, replicated the L1 version of Experiment I with a control group of native Dutch-speaking participants living in the Netherlands who were all 60 years or older. Although rhyme competition was not absent in the emigrants' L2, Experiment III nonetheless replicated the L2 version of Experiment I with native English-speaking participants in Australia, who were undergraduate students.

3 Experiment II

3.1 Participants

Twenty-two older adults, aged 62–85 years ($M = 69.8$, $SD = 6.5$; 11 females, 11 males), were recruited from the participant pool of the MPI for Psycholinguistics in Nijmegen, the Netherlands, and paid for their participation in this study. Results from seven additional participants were excluded due to calibration difficulties. All participants were native speakers of Dutch and had normal or corrected-to-normal vision. Written informed consent was obtained from each participant prior to the start of the experiment.
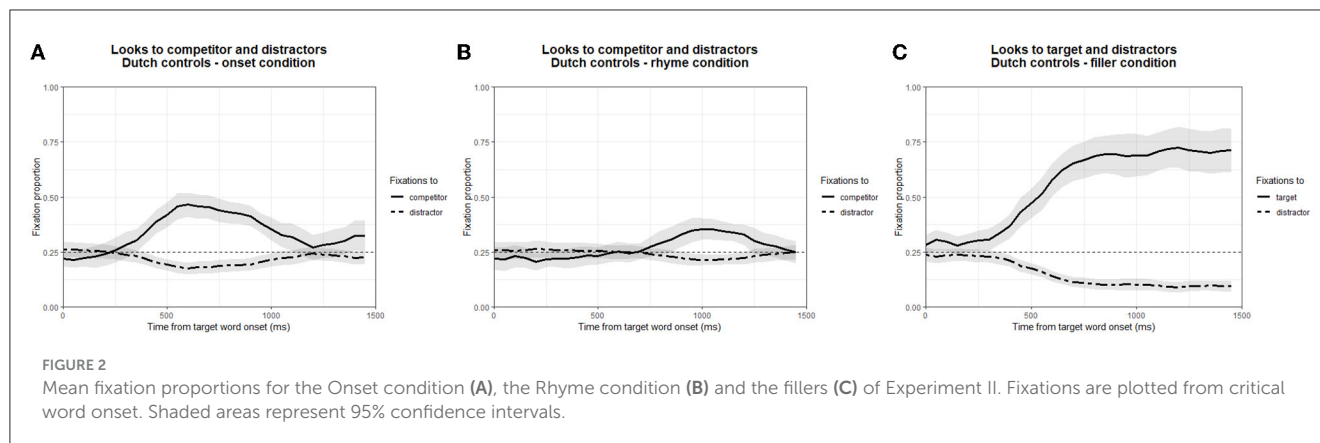
3.2 Materials and procedure

The stimulus materials were those from the L1 version of Experiment I. The procedure of the eyetracking task was the same as in the L1 session of Experiment I, with the exception that participants were seated in front of a computer screen at a viewing distance of 95 cm, and auditory stimuli were presented over Sennheiser HD201 headphones.

3.3 Results and discussion

As before, trials with a track loss percentage >30% were excluded from analysis. This was the case for 51 out of 1,320 trials. Mean fixation proportions to competitors and distractors (Onset and Rhyme condition) and to target and distractors (Filler trials) are shown in Figure 2, plotted from the onset of the critical word.

As in Experiment I, competitor fixation proportions in the one-second time window starting at 400 ms after critical word onset were compared to chance level (i.e., 0.25) using CPA. This showed that onset competitors were fixated significantly more than chance level (significant cluster from 400 to 1,140 ms; cluster summed t-value: 426.11; $p < 0.001$), as were rhyme competitors (significant cluster from 880 to 1,250 ms; cluster summed t-value: 137.20; $p < 0.001$). This finding indicates that, as expected, the native listeners of Dutch tested here experienced both onset and rhyme competition. This result thereby validates the stimuli and procedure of Experiment I. Since these listeners were all aged 60 and over, this replication furthermore confirms that the age of the emigrants who participated in Experiment I should not have precluded them from experiencing rhyme competition.

FIGURE 2
Mean fixation proportions for the Onset condition **(A)**, the Rhyme condition **(B)** and the fillers **(C)** of Experiment II. Fixations are plotted from critical word onset. Shaded areas represent 95% confidence intervals.

4 Experiment III

4.1 Participants

Twenty-five undergraduate students, aged 17–35 years ($M = 21.4$, $SD = 4.2$; 20 females, 5 males), from Western Sydney University participated in return for course credit. All participants were native speakers of Australian English with normal or corrected-to-normal vision. One further participant's data were excluded from analysis due to calibration difficulties. Each participant provided written informed consent before the start of the experiment.

4.2 Materials and procedure

The stimulus materials and procedure were those from the L2 version of Experiment I.

4.3 Results and discussion

Again, trials with a track loss percentage >30% were excluded from analysis (64 out of a total of 1,500 trials). Figure 3 shows the mean fixation proportions to competitors and distractors (Onset and Rhyme condition) and to target and distractors (Filler trials), plotted from the onset of the critical word.

As in both previous experiments, we conducted a CPA to compare competitor fixations in the time window from 400 to 1,400 ms to chance level (i.e., 0.25). This showed that both onset and rhyme competitors were fixated significantly above chance (Onset: significant cluster from 400 to 1,400 ms; cluster summed t-value: 406.05; $p < 0.001$; Rhyme: significant cluster from 930 to 1,270 ms; cluster summed t-value: 97.76; $p < 0.001$). This replication thus shows that native listeners of English experience both onset and rhyme competition. Again, the stimuli and procedure of Experiment I (here, L2 version) are validated.

5 General discussion

The processes by which spoken words are recognised can be reshaped by individual listeners' speech comprehension experience. In particular, the population of potential competitor words from among which a recognised word is chosen may be structured differently and become as a result either more or less comprehensive. While it is unsurprising that words with the same onset compete with one another for recognition, the further competition observed from word candidates with later overlap is less obviously useful in listening. Existing findings suggest that rhyme competition is modulable in that listeners may experience it more strongly when there is the possibility that there will be noise interruption to the speech signals (McQueen and Huettig, 2012; Brouwer and Bradlow, 2016). Our present results have further shown that rhyme competition can be shunned entirely.

The principal experiment in the present study compared the processes of within-language lexical activation and competition during L1 and L2 listening in a rare population of bilingual emigrants living in an L2 immersion environment. These processes proved to be largely the same in both languages, with no significant differences between languages in the time course of lexical activation, nor in the presence of onset competition. This is in line with the findings of Xue et al. (2020), who observed great similarity in the electrophysiological patterns produced as 34 young Mandarin-English L1-immersed bilinguals processed their two languages. As the L1 and L2 of the emigrants who participated in the present study (i.e., Dutch and English) are typologically much more similar than the languages spoken by Xue et al.'s bilinguals, our results suggests that phonological similarity vs. difference between a bilingual's languages is not a relevant factor in determining similarity or otherwise in how spoken words are processed. It would nevertheless be interesting to see further within-participant comparisons of L1 and L2 lexical activation and competition processes, in languages of varying typological similarity and by bilinguals of various kinds.

As expected, onset competition occurred in both of the emigrants' languages. Such competition is typically observed both in L1 listeners (e.g., Allopenna et al., 1998; Ben-David et al., 2011; McQueen and Huettig, 2012; Brouwer and Bradlow, 2016) and
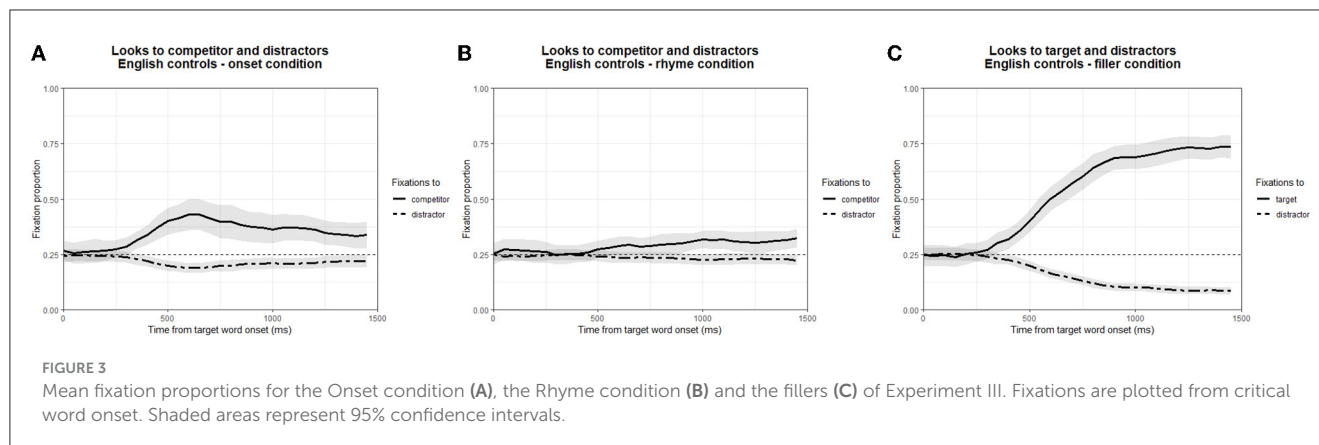
in L2 listeners (e.g., Mercier et al., 2014; Sarrett et al., 2021), and corresponds to the predictions of all models of spoken-word recognition (e.g., Cohort: Marslen-Wilson and Welsh, 1978; TRACE: McClelland and Elman, 1986; Shortlist: Norris, 1994; Shortlist B: Norris and McQueen, 2008). This finding suggests that the bilingual listeners were processing speech efficiently in both languages. That conclusion is fully supported by their looking patterns in the filler trials. In these trials—which indeed contained a target image on the screen—the bilinguals quickly recognised the spoken target word and looked at its referent picture. The present study thus revealed robust and efficient word recognition: the bilingual emigrants generally listen to speech and activate the necessary competitor population in each of their languages in the same way as the populations tested in all preceding studies in each language.

In the present study, rhyme competition was observed in the bilinguals' L2. The presence of this competition confirms the findings of the only other study to date that has examined rhyme competition in L2 listeners, which reported native-like rhyme competition effects for Korean L2 listeners of English (Shin et al., 2015). L2 listeners come in many flavours, of course, and the way they process language is always shaped by the type and amount of experience they have with each of their languages. The Korean-born listeners of English in the study by Shin and colleagues had been living in an English-speaking environment since on average 9.4 years of age, well within current assessments of the critical period for native language attainment (Hartshorne et al., 2018), so the findings of Shin et al. (2015) study may have been unsurprising. The participants in the present experiment, in contrast, were at least 18 years old (i.e., beyond the suggested threshold) when they took up residence in Australia. The finding that they nonetheless experienced L2 rhyme competition extends our existing knowledge about L2 rhyme competition to a new flavour of L2 listeners.

A perhaps somewhat surprising outcome of our study is the lack of rhyme competition in the bilinguals' L1. This contrasts with the rhyme competition that is typically attested in L1 listening (e.g., Allopenna et al., 1998; McQueen and Viebahn, 2007; McQueen and Huettig, 2012; Brouwer and Bradlow, 2016). What could be the reason for this finding? Rhyme competition effects are usually much weaker than onset competition, so could the modest number of emigrants tested in Experiment I have prevented us

from detecting rhyme competition? While we cannot rule it out definitively, this explanation seems unlikely since previous studies using the visual world paradigm have found rhyme competition with sample sizes of only 16 (McQueen and Viebahn, 2007) or even 12 (Allopenna et al., 1998) listeners. Our findings cannot be ascribed to the stimulus materials or experimental procedure either, as these did evoke rhyme competition in the same emigrants in the L2 version of Experiment I, as well as in both control groups of native listeners tested in Experiments II and III. Given that it should be presumed that the bilingual emigrants had previously experienced competition from rhyme competitors prior to migration, at least in the L1, we must therefore conclude that the absence of rhyme competition found here is either a general characteristic of L1 listening by fluent adult users of a late-acquired L2 (for whom no prior study of L1 rhyme competition has been, to our knowledge, undertaken), or it might be specific to our Experiment I participants. In the former case, it is clearly relevant that exactly such adult bilinguals show L2 effects on lexical processing in speech in the L1; for instance, Dutch-English bilinguals were slower recognising interlingual homophones (such as [li:f] which can be the Dutch word *lief* or the English word *leaf*) even in an experiment where English played no role at all (Lagrou et al., 2011), and similar unwanted onset competition from L2 in L1 listening has been demonstrated in both Spanish-English and German-English bilinguals (e.g., Blumenfeld and Marian, 2007, 2013). A multiplicity of onset (or whole-word) competitors uses up resources and may result in the non-availability of rhyme competitors ensuing by way of a trade-off.

On the other hand, the latter case—that the absence of rhyme competition may be true only for our particular participant population—could be due to the principal difference they showed from listeners in most previous studies: that they had left the L1 environment. After moving to Australia, these emigrants were no longer immersed in their L1 and instead used their L2, English, in most of their daily life. Recall that, when asked to specify their language use in a range of situations, our participants all indicated that they used their L2 more frequently than their L1 (see Appendix A in the Supplementary material). Most of what is known about L2 spoken-word recognition stems from research with less proficient L2 listeners in various stages of L2 learning. The emigrants in the current study, however, can be assumed to have

reached the end stage of learning: they are highly proficient users of the L2 for whom the L2 has become the dominant language. It is not inconceivable that this has influenced the processing of their L1. In fact, findings from a previous study we conducted with a group of Dutch emigrants that included most participants of the present study support this idea. In that study, we investigated these emigrants' ability to adapt to novel talkers in the L1 as well as the L2, using a lexically-guided perceptual learning paradigm (Bruggeman and Cutler, 2020). In this paradigm, participants are first exposed to an unusual sound in a lexical context that strongly favours a certain phonemic interpretation (e.g., an unusual [f] at the end of *behal-*). A subsequent phoneme categorisation task with speech from the same talker then reveals whether participants have adjusted the boundaries of their phoneme category for that talker's [f] to also include the unusual sound they were exposed to, indicating perceptual adaption to individual talkers. The group of Dutch emigrants tested by Bruggeman and Cutler (2020) showed robust adaptation in the L2, English, but no such adaptation in the L1, Dutch. The proposed explanation for this finding was that the emigrant participants no longer had as much need for the ability to accommodate for a speaker's atypical pronunciation in the L1, as they were rarely exposed to new, unfamiliar talkers in that language. This showed that mechanisms of L1 speech processing can decline—permanently or temporarily—after a long period of immersion in the L2 environment. The absence of L1 rhyme competition found in the present study extends the scope of this previous finding to a different level of processing.

The aforementioned explanations of the lack of rhyme competition in the L1—that it is typical for fluent adult users of a late-acquired L2, or that it specifically applies to the population tested here—are obviously testable in future studies of fluent adult bilinguals with a late-acquired L2, where the absence of rhyme competition in L1 may appear either in any environment at all, or only in an immersion L2 environment.

In sum, we conclude that the processes of lexical activation and lexical competition are largely the same in the L1 and the L2 for bilingual listeners who predominantly use their L2 in everyday life, and are highly proficient in their L2. Nevertheless, we further conclude that (a) listeners who live in an L2 immersion environment where they predominantly use the L2 may no longer experience rhyme competition in L1 listening, and (b) listeners with a late-acquired L2 who are highly proficient, indeed dominant, in their L2 can make use of rhyme competition as L1 listeners typically do.

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 Western Sydney University Human Research 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.

Author contributions

LB: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. AC: Conceptualization, Supervision, Writing – review & editing.

Funding

Acknowledgments

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/flang.2024.1275435/full#supplementary-material

References

Allopenna, P. D., Magnuson, J. S., and Tanenhaus, M. K. (1998). Tracking the time course of spoken word recognition using eye movements: evidence for continuous mapping models. *J. Mem. Lang.* 38, 419–439. doi: 10.1006/jmla.1997.2558

Apfelbaum, K. S., Blumstein, S. E., and McMurray, B. (2011). Semantic priming is affected by real-time phonological competition: evidence for continuous cascading systems. *Psychon. Bull. Rev.* 18, 141–149. doi: 10.3758/s13423-010-0039-8

Baayen, R. H., Piepenbrock, R., and Gulikers, L. (1995). *The Celex Lexical Database (CD-rom).* Philadelphia, PA: University of Pennsylvania, Linguistic Data Consortium.

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

Ben-David, B. M., Chambers, C. G., Daneman, M., Pichora-Fuller, M. K., Reingold, E. M., and Schneider, B. A. (2011). Effects of aging and noise on real-time spoken word recognition: evidence from eye movements. *J. Speech Lang. Hear. Res.* 54, 243–262. doi: 10.1044/1092-4388(2010/09-0233)

Blumenfeld, H. K., and Marian, V. (2007). Constraints on parallel activation in bilingual spoken language processing: examining proficiency and lexical status using eye-tracking. *Lang. Cogn. Process.* 22, 633–660. doi: 10.1080/01690960601000746

Blumenfeld, H. K., and Marian, V. (2013). Parallel language activation and cognitive control during spoken word recognition in bilinguals. *J. Cognit. Psychol.* 25, 547–567. doi: 10.1080/20445911.2013.812093

Boersma, P., and Weenink, D. (2013). *Praat: Doing Phonetics by Computer (Version 5.3.42).* Available online at: http://www.praat.org (accessed March 2, 2013).

Bregman, M. R., and Creel, S. C. (2014). Gradient language dominance affects talker learning. *Cognition* 130, 85–95. doi: 10.1016/j.cognition.2013.09.010

Broersma, M. (2012). Increased lexical activation and reduced competition in second-language listening. *Lang. Cogn. Process.* 27, 1205–1224. doi: 10.1080/01690965.2012.660170

Broersma, M., and Cutler, A. (2011). Competition dynamics of second-language listening. *Q. J. Exp. Psychol.* 64, 74–95. doi: 10.1080/17470218.2010.499174

Brouwer, S., and Bradlow, A. R. (2016). The temporal dynamics of spoken word recognition in adverse listening conditions. *J. Psycholinguist. Res.* 45, 1151–1160. doi: 10.1007/s10936-015-9396-9

Bruggeman, L., and Cutler, A. (2020). No L1 privilege in talker adaptation. *Biling.: Lang. Cogn.* 23, 681–693. doi: 10.1017/S1366728919000646

Canseco-Gonzalez, E., Brehm, L., Brick, C. A., Brown-Schmidt, S., Fischer, K., and Wagner, K. (2010). Carpet or cárcel: the effect of age of acquisition and language mode on bilingual lexical access. *Lang. Cogn. Process.* 25, 669–705. doi: 10.1080/01690960903474912

Chambers, C. G., and Cooke, H. (2009). Lexical competition during second-language listening: Sentence context, but not proficiency, constrains interference from the native lexicon. *J. Exp. Psychol.: Learn. Mem. Cogn.* 35, 1029–1040. doi: 10.1037/a0015901

Clopper, C. G., and Bradlow, A. R. (2009). Free classification of American English dialects by native and non-native listeners. *J. Phon.* 37, 436–451. doi: 10.1016/j.wocn.2009.07.004

Connine, C. M., Blasko, D. G., and Titone, D. (1993). Do the beginnings of spoken words have a special status in auditory word recognition? *J. Mem. Lang.* 32, 193–210. doi: 10.1006/jmla.1993.1011

Cutler, A., and Otake, T. (2004). Pseudo-homophony in non-native listening. Paper presented at the 147th meeting, Acoustical Society of America, New York, May 2004. *J. Acoust. Soc. Am.* 115, 2392. doi: 10.1121/1.4780547

Cutler, A., Weber, A., and Otake, T. (2006). Asymmetric mapping from phonetic to lexical representations in second-language listening. *J. Phon.* 34, 269–284. doi: 10.1016/j.wocn.2005.06.002

Díaz, B., Mitterer, H., Broersma, M., and Sebastián-Gallés, N. (2012). Individual differences in late bilinguals' L2 phonological processes: from acoustic-phonetic analysis to lexical access. *Learn. Individ. Differ.* 22, 680–689. doi: 10.1016/j.lindif.2012.05.005

Eisner, F., and McQueen, J. M. (2018). "Speech perception," in *The Stevens' Handbook of Experimental Psychology and Cognitive Neuroscience,* eds. S. Thompson-Schill (Hoboken: Wiley), 1–46.

Escudero, P., Hayes-Harb, R., and Mitterer, H. (2008). Novel second-language words and asymmetric lexical access. *J. Phon.* 36, 345–360. doi: 10.1016/j.wocn.2007.11.002

Farris-Trimble, A., and McMurray, B. (2013). Test-retest reliability of eye tracking in the visual world paradigm for the study of real-time spoken word recognition. *J. Speech Lang. Hear. Res.* 56, 1328–1345. doi: 10.1092-4388(2012/12-0145)

Garcia Lecumberri, M. L., Cooke, M., and Cutler, A. (2010). Non-native speech perception in adverse conditions: a review. *Speech Commun.* 52, 864–886. doi: 10.1016/j.specom.2010.08.014

Goodman, J. C., and Huttenlocher, J. (1988). Do we know how people identify spoken words? *J. Mem. Lang.* 27, 684–698. doi: 10.1016/0749-596X(88)90015-0

Hartshorne, J. K., Tenenbaum, J. B., and Pinker, S. (2018). A critical period for second language acquisition: evidence from 2/3 million English speakers. *Cognition* 177, 263–277. doi: 10.1016/j.cognition.2018.04.007

Holt, R., Bruggeman, L., and Demuth, K. (2021). Children with hearing loss can predict during sentence processing. *Cognition* 212, 104684. doi: 10.1016/j.cognition.2021.104684

Huettig, F., and Altmann, G. T. M. (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

Huettig, F., Rommers, J., and Meyer, A. S. (2011). Using the visual world paradigm to study language processing: a review and critical evaluation. *Acta Psychol.* 137, 151–171. doi: 10.1016/j.actpsy.2010.11.003

Ju, M., and Luce, P. A. (2004). Falling on sensitive ears: constraints on bilingual lexical activation. *Psychol. Sci.* 15, 314–318. doi: 10.1111/j.0956-7976.2004.00675.x

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

Lagrou, E., Hartsuiker, R. J., and Duyck, W. (2011). Knowledge of a second language influences auditory word recognition in the native language. *J. Exp. Psychol.: Learn. Memory Cognit.* 37, 952–965. doi: 10.1037/a0023217

Lagrou, E., Hartsuiker, R. J., and Duyck, W. (2013). Interlingual lexical competition in a spoken sentence context: evidence from the visual world paradigm. *Psychon. Bull. Rev.* 20, 963–972. doi: 10.3758/s13423-013-0405-4

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

Luce, P. A., Pisoni, D. B., and Goldinger, S. D. (1990). "Similarity neighborhoods of spoken words," in *Cognitive Models of Speech Processing: Psycholinguistic and Computational Perspectives*, eds. G. T. M. Altmann (Cambridge, MA: MIT), 122–147.

Luna, B., Velanova, K., and Geier, C. F. (2008). Development of eye-movement control. *Brain Cogn.* 68, 293–308. doi: 10.1016/j.bandc.2008.08.019

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

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

Marslen-Wilson, W. D., and Welsh, A. (1978). Processing interactions and lexical access during word recognition in continuous speech. *Cogn. Psychol.* 10, 29–63. doi: 10.1016/0010-0285(78)90018-X

Matin, E., Shao, K. C., and Boff, K. R. (1993). Saccadic overhead: Information-processing time with and without saccades. *Perception and Psychophysics* 53, 372–380. doi: 10.3758/BF03206780

MATLAB (Version R2019b) [Computer software]. Natick, MA, United States: The MathWorks, Inc.

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

McMurray, B. (2017). *Nonlinear Curvefitting for Psycholinguistics (Version 24).* Available online at: https://osf.io/4atgv/ (accessed November 30, 2020).

McQueen, J. M., and Huettig, F. (2012). Changing only the probability that spoken words will be distorted changes how they are recognized. *J. Acoust. Soc. Am.* 131, 509–517. doi: 10.1121/1.3664087

McQueen, J. M., and Viebahn, M. C. (2007). Tracking recognition of spoken words by tracking looks to printed words. *Q. J. Exp. Psychol.* 60, 661–671. doi: 10.1080/17470210601183890

Mercier, J., Pivneva, I., and Titone, D. (2014). Individual differences in inhibitory control relate to bilingual spoken word processing. *Biling.: Lang. Cogn.* 17, 89–117. doi: 10.1017/S1366728913000084

Miller, J., Patterson, T., and Ulrich, R. (1998). Jackknife-based method for measuring LRP onset latency differences. *Psychophysiology* 35, 99–115. doi: 10.1111/1469-8986.3510099

Norris, D. (1994). Shortlist: a connectionist model of continuous speech recognition. *Cognition* 52, 189–234. doi: 10.1016/0010-0277(94)90043-4

Norris, D., and McQueen, J. M. (2008). Shortlist B: a Bayesian model of continuous speech recognition. *Psychol. Rev.* 115, 357–395. doi: 10.1037/0033-295X.115.2.357

Norris, D., McQueen, J. M., and Cutler, A. (1995). Competition and segmentation in spoken-word recognition. *J. Exp. Psychol: Learn. Memory Cognit.* 21, 1209–1228. doi:;10.1037/0278-7393.21.5.1209

R Core Team (2019). *R: A Language and Environment for Statistical Computing (Version 3.6.1)*. Vienna, Austria: R Foundation for Statistical Computing. Available online at: https://www.R-project.org/ (accessed October 3, 2022).

Sarrett, M. E., Shea, C., and McMurray, B. (2021). Within- and between-language competition in adult second language learners: implications for language proficiency. *Lang. Cognit. Neurosci.* 37, 165–181. doi: 10.1080/23273798.2021.1952283

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

Shin, H. A., Bauman, B., MacPhee, I. X., and Zevin, J. D. (2015). "The dynamics of spoken word recognition in second language listeners reveal native-like lexical processing," in *Proceedings of the 37th Annual Meeting of the Cognitive Science Society*, eds. D. C. Noelle, R. Dale, A. S. Warlaumont, J. Yoshimi, T. Matlock, C. D. Jennings, and P. P. Maglio (Austin: Tx), 2158–2163.

Slowiaczek, L. M., Nusbaum, H. C., and Pisoni, D. B. (1987). Phonological priming in auditory word recognition. *J. Exp. Psychol.: Learn. Memory Cognit.* 13, 64–75. doi: 10.1037/0278-7393.13.1.64

Smulders, F. T. Y. (2010). Simplifying jackknifing of ERPs and getting more out of it: retrieving estimates of participants' latencies. *Psychophysiology* 47, 387–392. doi: 10.1111/j.1469-8986.2009.00934.x

Spivey, M. J., and Marian, V. (1999). Cross talk between native and second languages: partial activation of an irrelevant lexicon. *Psychol. Sci.* 10, 281–284. doi: 10.2307/40063426

Taft, M., and Hambly, G. (1986). Exploring the cohort model of spoken word recognition. *Cognition* 22, 259–282. doi: 10.1016/0010-0277(86)90017-X

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., and Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science* 268, 1632–1634. doi: 10.1126/science.7777863

Ulrich, R., and Miller, J. (2001). Using the jackknife-based scoring method for measuring LRP onset effects in factorial designs. *Psychophysiology* 38, 816–827. doi: 10.1111/1469-8986.3850816

Weber, A., and Cutler, A. (2004). Lexical competition in non-native spoken-word recognition. *J. Mem. Lang.* 50, 1–25. doi: 10.1016/S0749-596X(03)00105-0

Xue, J., Li, B., Yan, R., Gruen, J. R., Feng, T., Joanisse, M. F., et al. (2020). The temporal dynamics of first and second language processing: ERPs to spoken words in Mandarin-English bilinguals. *Neuropsychologia* 146, 107562. doi: 10.1016/j.neuropsychologia.2020.107562