Lexical competition in non-native spoken-word recognition

Andrea Weber[a,*] and Anne Cutler[b]

[a] *University of the Saarland, Saarbrücken, Germany*
[b] *Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands*

Abstract

Four eye-tracking experiments examined lexical competition in non-native spoken-word recognition. Dutch listeners hearing English fixated longer on distractor pictures with names containing vowels that Dutch listeners are likely to confuse with vowels in a target picture name (*pencil,* given target *panda*) than on less confusable distractors (*beetle,* given target *bottle*). English listeners showed no such viewing time difference. The confusability was asymmetric: given *pencil* as target, *panda* did not distract more than distinct competitors. Distractors with *Dutch* names phonologically related to English target names (*deksel*, 'lid,' given target *desk*) also received longer fixations than distractors with phonologically unrelated names. Again, English listeners showed no differential effect. With the materials translated into Dutch, Dutch listeners showed no activation of the English words (*desk*, given target *deksel*). The results motivate two conclusions: native phonemic categories capture second-language input even when stored representations maintain a second-language distinction; and lexical competition is greater for non-native than for native listeners.
© 2003 Elsevier Inc. All rights reserved.

Keywords: Spoken-word recognition; Non-native listening; Lexical competition; Phonetic contrast

Listening to one's native language is effortless; but listening to a second language can be distressingly hard work. Unfamiliar words, unknown idioms, and hitherto unencountered accents can at any moment present new challenges. Speech can seem unnervingly fast, because procedures for segmenting speech of the native language into words fail to work with the second language. The experience of being tired out by simply listening, for instance to a lecture or a theatre performance, is one that many second-language listeners have undergone.

In one respect, however, there could be a ray of hope for the non-native listener. The recognition of spoken words is, after all, based on a process of multiple simultaneous activation of word candidates (e.g., Marslen-Wilson & Welsh, 1978; Zwitserlood, 1989) and subsequent competition between them (e.g., Goldinger, Luce, Pisoni, & Marcario, 1992; McQueen, Norris, & Cutler, 1994).

The more active candidates and the more competition, the slower recognition proceeds (e.g., Norris, McQueen, & Cutler, 1995; Vroomen & de Gelder, 1995). Thus, having a smaller vocabulary should at least benefit the recognition of the members of that small set if they are presented, simply because the remainder of the vocabulary will offer less interfering competition. The word *balance*, for instance, might be subject to initial competition from *ballad*, *ballast*, *ballot*, *balustrade*, and *balneology* for a speaker with a very well-stocked vocabulary; but the fewer of those lower-frequency competitors that are activated in a listener's individual vocabulary, the less competition will be offered when *balance* occurs in conversation. Non-native listeners are likely to know fewer such words than native listeners. It is thus at least conceivable that in this small respect non-native listening could have, at least when known words are presented, a small advantage over native listening.

The processes involved in spoken-word recognition are presumably universal. Thus we assume that there

---

* Corresponding author. Fax: +49-681-302-6561.
  E-mail address: aweber@coli.uni-sb.de (A. Weber).

will be competition during spoken-word recognition in a non-native language just as in the native language. The question at issue is thus not whether such competition occurs, but how much. Is the extent of competition in second-language listening noticeably limited, in comparison to native listening? Is vocabulary size in the target language the sole relevant determinant of the amount of competition?

Recent evidence suggests that it is not. Unfortunately for the non-native listener, there appear to be several factors at work that can act to *increase* competition in second- by comparison with first-language processing. A major complicating factor is that speakers of more than one language seem not fully able to keep language vocabularies apart. Studies from visual word recognition certainly suggest that first- and second-language vocabularies can both be activated together by written words; thus written forms that are cross-language homographs such as BRAND (English *brand*, Dutch 'fire') are responded to slower in visual lexical decision by participants who know both the languages in question (Dijkstra, van Jaarsveld, & ten Brinke, 1998), and can elicit false negative responses when the task is to pick out the English words from a mixed list of English and Dutch words (Dijkstra, Timmermans, & Schriefers, 2000). Similarly, masked priming from very brief presentations of letter strings can activate words in more than one vocabulary even when participants think they are reading words in only one language (Bijeljac-Babic, Biardeau, & Grainger, 1997). Competition from the rest of the vocabulary is correspondingly increased since not only are homographic forms available from the interfering language, but also similar words and partially overlapping forms; thus the recognition of visually presented English words by Dutch readers is affected by the number of similarly spelled words in the Dutch vocabulary (van Heuven, Dijkstra, & Grainger, 1997). These latter studies have in common that they showed the influence of the native lexicon on the processing of a non-native language without actually presenting a stimulus of the native language. They demonstrated that bilinguals could not deactivate the lexicon of the native language in these experiments even in a monolingual non-native situation where the native vocabulary is irrelevant.

The graphemes used to represent language visually are of course the same in a given orthography irrespective of which language it is representing: English BRAND and Dutch BRAND look alike. This is not true of the phonemes that constitute speech information; only rarely do the realizations of phonemic categories map exactly onto those of another language. Thus, one might argue that simultaneous activation of the vocabularies of two languages might be avoidable in listening, since acoustic–phonetic information would cue the listener as to which language should be active. Certainly

there is evidence from gating that listeners can detect the language in which a word is spoken on the basis of very little acoustic information, even if the language of the surrounding spoken context is different (Grosjean, 1988; Li, 1996).

Nevertheless, spoken-word recognition by listeners with more than one language also appears to be bedeviled by multiple vocabulary activation. Indirect evidence again comes from visual lexical-decision studies, which have shown that word recognition in a non-native language is sensitive not only to visual but also to phonological cross-language similarity. Thus, Dutch–English bilinguals were slower to reject non-words that were cross-language pseudohomophones (e.g., the English non-word *SNAY*, which, pronounced according to English spelling-to-sound conversion rules, sounds like the Dutch word *snee*, 'slice') than regular non-words (e.g., *ROLM*; Nas, 1983). Doctor and Klein (1992) also found inhibitory effects of cross-language homophony for English–Afrikaans bilinguals, and Dijkstra, Grainger, and van Heuven (1999) found inhibitory effects of phonological cross-language overlap for Dutch–English bilinguals. Direct evidence from spoken-word processing was provided by Schulpen, Dijkstra, Schriefers, and Hasper (in press), who found that recognition of visually presented LEASE was facilitated by prior auditory presentation of the effectively homophonic Dutch *lies*, 'groin,' even though variation in the production of /liːs/ across the Dutch and English words, respectively, was detected by listeners in a gating task. Spivey and Marian (1999; see also Marian & Spivey, 1999, 2003, in press) also used a spoken-word recognition task, showing via measurement of eye movements that words of two languages were activated even though words of only one language were heard.

Moreover, the lack of overlap in the phonemic repertoires of two languages may not actually act to reduce cross-language competition at all. It may, instead, increase it. This is because non-native phoneme perception is often inaccurate, as has been amply attested in speech perception studies (for an overview see Strange, 1995). Phonemic categories of a second language learned after childhood are notoriously hard to acquire, and discrimination of the contrasts involved may never reach native standards. Most importantly, the categorical distinctions of the first language prevail, so that the hardest second-language contrasts to learn are those which are ignored in the native language because each of the contrasting sounds is a permissible token of a single native category (Best, 1995). Japanese listeners, for example, have notorious difficulty in distinguishing English /r/ and /l/, which both map (badly) to a single Japanese category.

The obvious effect of this is that minimal pairs involving the contrast in question will not be easy for second-language listeners to tell apart. Japanese listeners

may thus perceive *right* and *light* as homophones to the same extent as *sale* and *sail*, or at the very least they will perceive *right* and *light* as much more similar than they will sound to native listeners. And this, in turn, will increase competition in that the potential set of word candidates activated by presentation of any second-language word containing one of the confusable phonemes could involve activation also of words containing the other. Thus, the beginning of *balance* may activate not only words beginning *bal-* but also those beginning *bar-*, such as *barren*, *barrow*, and *barrister*.

Still more problematically, the capture of second-language phonemic input by first-language phonemic categories could actually exacerbate the availability of the native vocabulary; not only near-homophones, but spurious competitors based on native phonemic mappings could become activated. By this token *balance* could also activate a Japanese listener's native *bara*, 'rose,' or *baratsuku*, 'vary.'

In short, a good part of the notorious difficulty of listening to speech in a non-native language could arise because of unwanted activation of spurious competitor words. Even though the non-native listener knows fewer words of the language than the native listener, the total competitor population in any given word recognition operation may be larger, because inaccurate phonetic processing allows spurious candidates from the native language on the one hand and spurious phonemic matches in the second language on the other. In the present study, this is the issue we address.

Our experiments use a methodology eminently suited to the investigation of competitor activation, namely listening plus the recording of gaze via a head-mounted eye-tracker (Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995; for an overview of the paradigm see Tanenhaus & Spivey-Knowlton, 1996). This method exploits the fact that participants make saccadic eye movements to either real objects or pictures of objects on a screen as the names of the objects are mentioned in spoken instructions. Locations and latencies of eye movements to pictures can therefore be used to examine lexical access in spoken-word recognition. Since eye movements can be continuously recorded, it is possible to monitor the comprehension process as spoken language unfolds over time, and hence evaluate relative competitor activation over time as well.

Tanenhaus et al. (1995) presented American English speaking participants with a display of objects that sometimes included two objects with initially similar names (e.g., *candy* and *candle*) and instructed them to move the objects around on a table. They found that the mean time to initiate an eye movement to the correct object (e.g., *candy*) was longer when the display included an object with a phonologically similar name (e.g., *candle*) than when no such object appeared. Later studies replicated this competition effect. For French,

Dahan, Swingley, Tanenhaus, and Magnuson (2000) showed that participants fixated pictures with names sharing initial sounds with the target (e.g., *bouteilles*, 'bottles' given target *boutons*, 'buttons') more than pictures with phonologically unrelated names. For Dutch, Salverda, Dahan, and McQueen (in press) found competition for fixation between a pictured *blik*, 'can,' and *bliksem*, 'lightning.' Allopenna, Magnuson, and Tanenhaus (1998) found clear activation not only for competitors with names that share the onset with a target (e.g., target *beaker* and onset competitor *beetle*) but also for competitors that share the rhyme with a target (e.g., target *beaker* and rhyme competitor *speaker*). Both Allopenna et al. (1998) and Dahan, Magnuson, and Tanenhaus (2001) compared fixation proportions with fixation probabilities derived from activations over time for targets and competitors in simulations with the TRACE model (McClelland & Elman, 1986), and found very similar shapes of functions. The results suggest that competition effects defined as fixation proportions to pictures can indeed be closely mapped to activation levels of word candidates over time. This is seen as indicating that the observed competition effects are not caused by circumscribed visual contexts, but indeed represent natural language processing (for a discussion of this issue see, Tanenhaus, Magnuson, Dahan, & Chambers, 2000).

In the non-native listening study of Spivey and Marian (1999), referred to above, eye movements of Russian–English bilinguals to objects that were displayed on a table were monitored. The participants lived in the US, and English had been their primary language for an average of four years. Participants differed with respect to their self-reported language preferences at the time of the study. In separate sessions, participants were instructed in Russian and English to move objects on a table. In one condition in the Russian session, the target object was accompanied by an object whose English name shared initial sounds with the Russian target (e.g., Russian target *marku*, 'stamp,' and English competitor *marker*). In another condition, the cross-language competitor was replaced by an unrelated distractor whose name bore no similarity to the target (e.g., Russian *lineika*, 'ruler'). Similarly, in the English session the English target object was either accompanied by a Russian competitor or not. Across the two sessions, participants made more eye movements to the cross-language competitors than to the unrelated distractors on average. However, when the sessions were analyzed separately, significant competition from English items during the Russian session was found, but no significant competition from Russian items during the English session. In a follow-up study, Marian and Spivey (1999) found the opposite pattern: competition from Russian for English, but no competition from English for Russian. They explained this asymmetry by reference to

manipulations of the language mode during the experiment and general language preferences of the participants. In the Marian and Spivey (1999) study more effort was made to put the participants into a Russian language mode by giving instructions in Russian and additionally playing popular Russian songs at the beginning of the Russian session. Marian and Spivey's analyses did not address the precise relationship between the phoneme inventories of Russian and English.

In the present study, we focused first on establishing whether spurious phonemic matches in the second language would lead to increased competition. The phonemic contrasts we used were English vowel contrasts that are attested to be difficult for Dutch listeners to apprehend. The non-native participants were native Dutch listeners, highly proficient in English. They had all learned English as a second language in school, and Dutch was clearly their primary language; however, their English competence was also very high, sufficient for instance to follow lectures in English, since that requirement is considered normal for undergraduates at Dutch universities. The experiments were conducted completely in English, and participants were not aware of the fact that their native language was relevant. Nonetheless, no long-term shifts in language mode were expected, since the participants lived in the Netherlands at the time of the study (cf. Flege & Frieda, 1997; Meador, Flege, & MacKay, 2000).

The participants were instructed in spoken English to click on pictures of objects on a computer screen. A target picture (e.g., of a panda) was one of four pictures. The name of one distractor picture shared initial segments with the name of the target picture (for example, target *panda*, competitor *pencil*). Half of the target–competitor pairs contained English vowels often confused by Dutch listeners (e.g., /æ/ and /ɛ/ as in *panda pencil*), the other half contained vowels that are unlikely to be confused (e.g., /ɒ/ and /i/ as in *bottle–beetle*). If phonetic discrimination difficulties affect non-native listeners' competition, then Dutch listeners should fixate distractor pictures with confusable English vowels longer than distractor pictures with distinct vowels.

Although all vowel contrasts of a language must be able to distinguish one word from another for native listeners (otherwise they would not be contrasts), it is for instance true that the vowels /æ/ and /ɛ/ are closer in acoustic space than are the vowels /ɒ/ and /i/, with the result that the former pair is more likely than the latter pair to be confused by native listeners in perceptual identification (Hillenbrand, Getty, Clark, & Wheeler, 1995) and in word recognition (Broersma, 2002). As a control for effects of inherent phonetic confusability, independent of the abilities of the listeners to resolve the confusion, we therefore also presented the same stimuli to listeners whose native language was English. If the Dutch listeners show the predicted asymmetry of

response for the two types of contrast, but this is simply due to inherent confusability, the same results will appear with native listeners. But if native listeners can indeed discriminate the vowel contrasts of their language rapidly and easily within a real-word context, then we expect a different pattern of results. Following brief competitor activation in both target–competitor pairs (irrespective of vowel confusability, and merely resulting from the shared initial consonant of target and competitor), activation should decrease at the same rate for competitors with confusable and with distinct vowels as soon as acoustic information about this distinguishing vowel becomes available.

Experiment 1

Method

Participants

Twenty University of Nijmegen students took part in Experiment 1a. They were native speakers of Dutch who had lived in the Netherlands all their lives, and had normal or corrected-to-normal vision and normal hearing. They had received an average of 7.8 years of training in English as a foreign language in secondary education beginning at a mean age of 11. Twelve native speakers of British English participated in Experiment 1b. Most of these were students, and at the time of testing had been in the Netherlands for no more than four weeks. They had at best rudimentary knowledge of Dutch. All participants received a small payment for taking part.

The Dutch participants underwent a multiple-choice test in English after completing the eye-tracking experiment to confirm their high proficiency in the non-native language. For 20 nouns (none of which occurred in the eye-tracking experiment), they had to choose the correct English definition out of three possibilities. The definitions for the nouns were taken from the Longman Dictionary of Contemporary English (1987). Most false definitions described nouns that were either phonologically or semantically related to the target noun (e.g., the definition for *brunch* was an option for the phonologically related target word *branch*, the definition for *fountain* was an option for the semantically related target word *river*). The average score was 97% correct.

Materials

Twenty English nouns referring to picturable objects (e.g., *panda*) were chosen as target words. Each target word was paired with a competitor. The onset of the competitor overlapped phonemically with the onset of the target word. Ten target–competitor pairs contained in the first syllable vowels which previous research (Broersma, 2002; Cutler, Weber, Smits, & Cooper, forthcoming) had shown to be confusable for Dutch

listeners. In these cases the initial CVC of target and competitor putatively overlapped. We selected the two vowel pairs in question on the basis of phoneme identification data from an experiment in which Dutch listeners had categorized American English phonemes (Cutler et al., forthcoming); they were the lax vowel pair /æ/-/ɛ/ and the diphthong pair /aɪ/-/eɪ/. Dutch contains a vowel labeled /ɛ/ and a diphthong labeled /ɛɪ/, but no /æ/ or /aɪ/. In both cases, the identification results had shown an asymmetry, one vowel being confused for the other more than vice versa.

Target words with the vowels /æ/ or /aɪ/ in the first syllable were paired with competitors containing the vowels /ɛ/ or /eɪ/, respectively, (e.g., target *panda* /pændə/ and competitor *pencil* /pɛnsl/; target *tie* /taɪ/ and competitor *tail* /teɪl/). The other 10 target–competitor pairs contained each a different combination of vowels, some spectrally closer, some further apart; again, the same English phoneme identification data from Dutch listeners (Cutler et al., forthcoming) had shown that these pairs were not confusing for Dutch listeners. Examples are target *bottle* /bɒtl/ and competitor *beetle* /bitl/. (In those cases overlap comprised initial C_C). Bearing in mind the Marian and Spivey (1999) finding of competition from the native lexicon, we also ensured that whenever a target word's Dutch translation began in the same way as the target word, the relevant competitor's Dutch translation also began in the same way as the competitor (e.g., English target *panda* /pædə/ is *panda* /panda/ in Dutch and English competitor *pencil* is *potlood* /pɔtlot/ in Dutch).

Two phonologically unrelated distractors were added for each target word (e.g., *strawberry* and *dress*). Neither the English nor the Dutch names of the unrelated distractors (e.g., *aardbei* /ardbɛɪ/, 'strawberry' /strɔːbərɪ/ and *jurk* /jʉrk/, 'dress' /drɛs/) overlapped with the English target word. The pictures of a target item, its competitor, and two unrelated distractors were displayed together in one trial set. The target word was actively named in the spoken instructions, whereas the competitor and the unrelated distractors were not named. The 20 target words, their competitors, and unrelated distractors are listed in Appendix A. To prevent participants developing expectations that pictures with phonologically similar names were likely targets, 20 additional filler trials were constructed, again with four items each (e.g., *candy*, *pig*, *ashtray*, *dice*). In the filler trials, no phonemic overlap occurred between the Dutch or English names of the items. For example, for the picture of a *candy* as target, neither the English nor Dutch names of the other three items in that trial had initial /k/. Finally, six representative trials were also constructed as practice trials.

The pictures of the items were selected from the Snodgrass and Vanderwart (1980) and the Cycowicz, Friedman, Rothstein, and Snodgrass (1997) picture sets, as well as from the Art Explosion library (1995). All pictures were black and white line drawings. In order to establish naming norms, 10 native speakers of Dutch were asked to name and rate the targets' and their competitors' pictures in English. The agreement between participants' responses and the intended names was 95% and the goodness was rated with a mean of 4.9 on a scale from 0 to 7. Some small suggestions for improvement of the pictures were implemented. None of the participants from the naming and rating experiment took part in the eye-tracking experiment.

Previous research has shown that high-frequency candidate picture names (e.g., *window*) are more likely to lead to fixations than low-frequency candidates (e.g., *windmill*; Dahan et al., 2001). For the present study the English word form frequencies of the targets (mean 31.28 per million) and the competitors (30.14 per million) were computed using the *CELEX* database (Baayen, Piepenbrock, & van Rijn, 1993). No significant main effects or interactions were observed in a mixed two-factor ANOVA with picture (with the two levels 'target' and 'competitor') as within-items factor and vowel pair (with the two levels 'confusable' and 'distinct') as between-items factor.[1] Dahan et al. (2001) also showed that the probability of fixating distractors that do not match the acoustic information of the target word (e.g., *turkey* or *pipe*, given input beginning *win-*) does not vary with lexical frequency. Therefore, in the present study lexical frequencies of the unrelated distractors were not controlled.

The spoken instructions were recorded onto DAT in a soundproof booth by a male native speaker of British English, sampling at 48 kHz. The material was then down-sampled to 16 kHz and stored on disc. Durations of the preceding contexts and the target words were measured using the Xwaves software. An instruction was, for example: *Click on the panda. Now put it on top of the circle*. The average duration of the preceding context (*click on the*) was 392 ms, of the target word in confusable target–competitor pairs (e.g., *panda*) 548 ms, and of target word in distinct pairs (e.g., *bottle*) 516 ms. In addition, the duration of the putative overlap between the target word and its competitor (e.g., the duration of /pæn/ in *panda*) was measured. The average duration of this portion was 292 ms for confusable target–competitor pairs and 386 ms for the equivalent portion of distinct pairs (e.g., the duration of /bɒt/ in *bottle*).

Procedure

The auditory stimuli were presented over headphones using the NESU experiment control software

---

[1] Equivalent analyses with English and Dutch lemma frequencies (the sum of the appropriate word form frequencies of a lexical entry, e.g., singular plus plural word form frequencies of a noun) revealed similar results.

(see http://www.mpi.nl/tools/nesu.html). Participants' eye movements were monitored using a SMI EyeLink-Hispeed 2D eye-tracking system. Two cameras on a lightweight headband provided the input to the tracker. The center of the pupil was tracked to determine the position of the eye relative to the head. Throughout the experiment, the computer recorded the onset and offset times and the spatial coordinates of the participants' fixations. The signal from the eye tracker was sampled every 4 ms. Both eyes were monitored, but we analyzed only data from the right eye. Along with the eye movements, the position of the mouse cursor was recorded while participants manipulated objects on the computer screen.

Participants were tested individually. At the beginning of a session they received written instructions in English, that included an example of a trial display and an explanation of the task. Participants were then seated in a quiet room approximately 60 cm in front of a monitor. After the eye tracker was calibrated, each participant was presented with the 46 trials (6 practice trials plus 20 experimental trials plus 20 filler trials). Each experimental trial was preceded by a filler trial. All pictures were presented as white line drawings on a blue background on a 5×5 gray grid. In each trial, four line-drawing pictures and four green geometric shapes, each scaled to fit into a cell of the grid, and a cross centered in the middle, appeared on the screen (see Fig. 1). Each cell measured 4.3×4.3 cm, corresponding to a visual angle of approximately 4°, which is well within the resolution of the tracker (0.1°). The positions of the target object and its competitor were randomized across trials. The positions of the geometric shapes were fixed, and participants were told this in advance. Spoken instructions started simultaneously with the appearance of the pictures on the screen. Participants were first asked to click on one of the four pictures using the mouse (e.g., *Click on the panda.*), and then to move the picture on top of one of the four geometric shapes (e.g., *Now put it on top of the circle.*). Once this was accomplished, the experimenter initiated the next trial.

Following Dahan et al. (2000, 2001), the set of pictures was not shown to the participants before the experiment. There was also no delay between the appearance of the display on the screen and the beginning of the spoken instructions. This procedure makes it less likely that participants have implicitly named the pictures beforehand. There was also no instruction to fixate the cross at the onset of the trials. Therefore, participants could be fixating any of the four objects or the cross at the onset of the target word. (Only very rarely did participants fixate any other location on the screen.)

After every five trials a centered fixation point appeared, and participants were instructed to look at it. The experimenter could then correct potential drifts in the calibration of the eye tracker. The experiment lasted approximately 10 min.

Graphical software was used to display the locations of the participants' fixations as dots superimposed on the four line drawings for each trial and each participant. The fixation dots were numbered in the order in which the fixation had been produced. Onset times and durations of the fixations were displayed in another window. Fixations on the line drawings were coded as pertaining to the cell of the target object, the competitor, or one of the two unrelated distractors. Fixations that lay clearly outside the cell of an object were not used for the computation of fixation proportions. For each trial, fixations were coded from target word onset until the participant had clicked with the mouse cursor on the target picture, which was taken as the participant's identification of the target. Saccade times were not added to fixation times.

Results and discussion

Experiment 1a

Eight trials were removed from the analysis because participants clicked on an object other than the target object without correcting their choice (2% of all trials). This low percentage of errors suggests that Dutch participants had no problem performing the task. Fixation proportions were averaged over participants and items for separate analyses. Blinks and saccades were not included in the calculation of the fixation proportions.

Fig. 2 presents the proportions of fixations averaged over participants for trials with confusable vowels in the target and its competitor (Fig. 2a) and for trials with distinct vowels (Fig. 2b). Fixation proportions for the
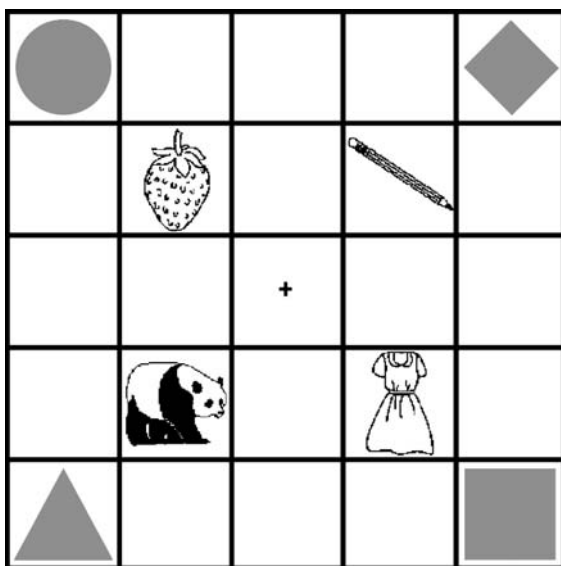

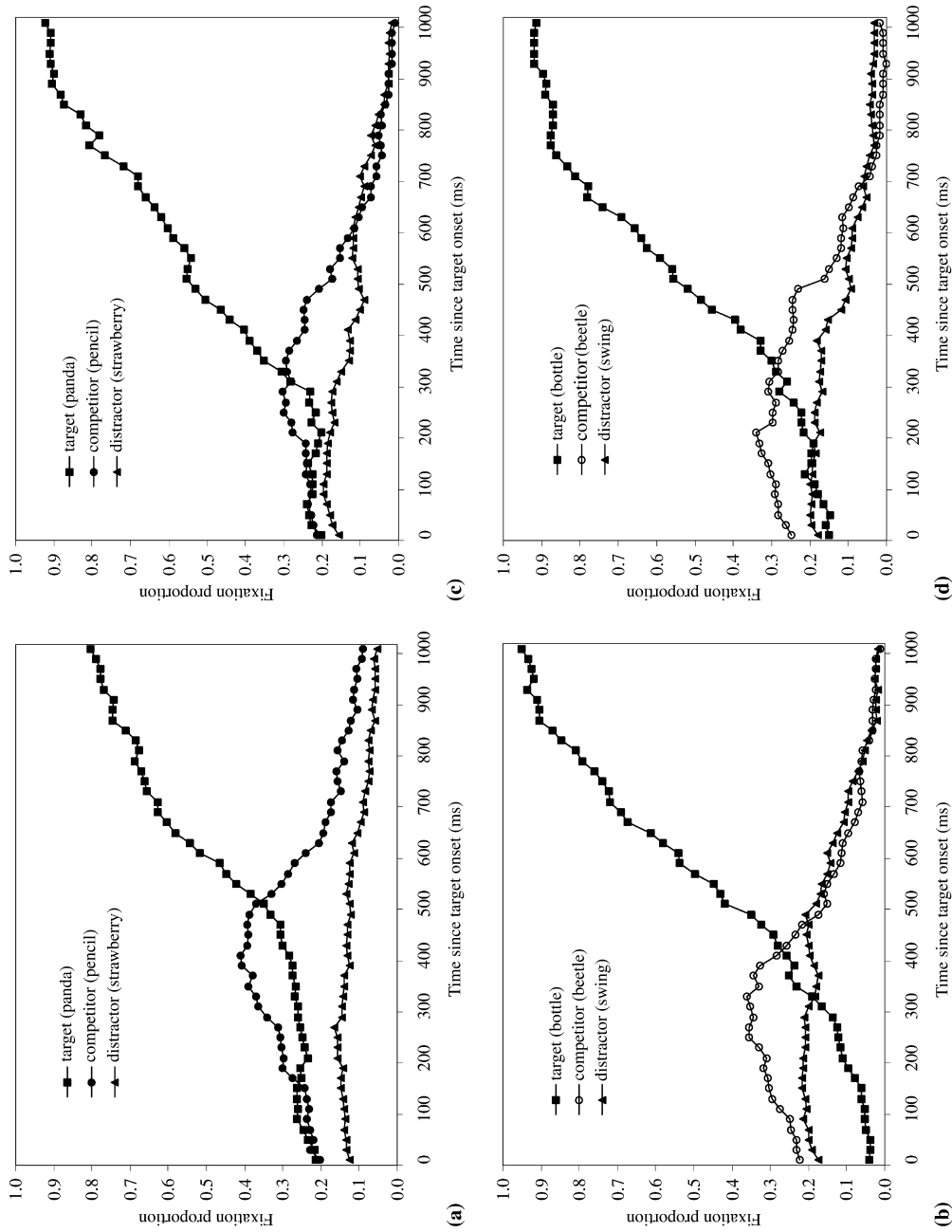Fig. 1. Example of a visual display presented to participants.

Fig. 2. Experiment 1. Fixation proportions over time for English targets, competitors, and averaged distractors: (a) Dutch listeners, target vowel confusable for Dutch listeners with competitor vowel; (b) Dutch listeners, target vowel distinctive for Dutch listeners from competitor vowel; (c) British English listeners, target vowel confusable for Dutch listeners with competitor vowel; and (d) British English listeners, target vowel distinctive for Dutch listeners from competitor vowel.
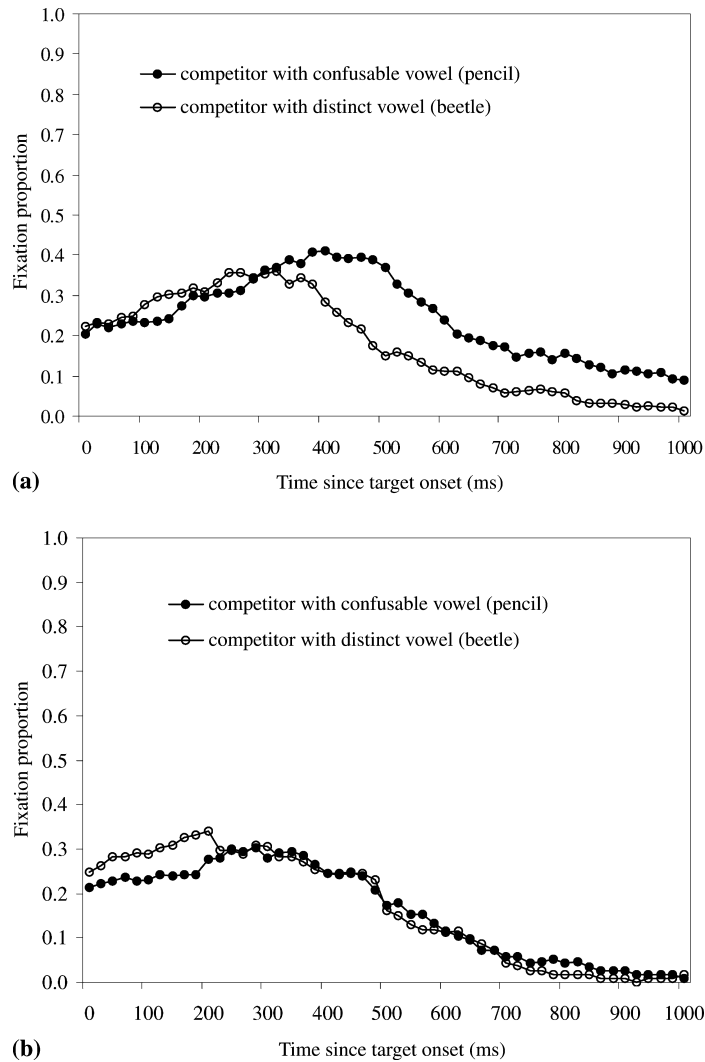
Fig. 3. Experiment 1. Fixation proportions over time for English competitors when target vowel was confusable (Figs. 2a and c) and when target vowel was distinctive (Figs. 2b and c): (a) Dutch listeners and (b) British English listeners.

two unrelated distractors were averaged. In Fig. 3a, the proportions of fixations for the two competitors from Fig. 2 are displayed. All figures show fixation proportions in 20 ms time slices from 0 to 1000 ms after target word onset. It is estimated that an eye movement is typically programmed about 200 ms before it is launched (e.g., Fischer, 1992; Matin, Shao, & Boff, 1993; Saslow, 1967), so that 300 ms after target onset is approximately the point at which fixations driven by the first 100 ms of acoustic information from the target word can be seen. Thus, the mapping of the acoustic signal onto the lexical representations is reflected by fixations from about 300 ms on. Just before or at this point, fixation proportions to the targets should begin to increase, and

fixation proportions to unrelated distractors should begin to decrease (see for example, Allopenna et al., 1998; Dahan et al., 2000).

As can be seen in Fig. 3a, the listeners fixated competitors with confusable vowels more often and longer than competitors with distinct vowels. Around 300 ms the fixation proportions to the two competitors started to diverge, with fixation proportions to confusable competitors remaining higher, at least until 800 ms after word onset. Since the fixation proportions in the time window from 0 to 300 ms were equally high for the two competitors (F_1 and $F_2 < 1$), a direct comparison of the further course of fixation proportions seemed justified. Over the 300–800 ms time window fixa-

tion proportions were 28.7% to competitors with confusable vowels and 17% to competitors with distinct vowels.[2] A one-factor ANOVA on the mean proportion of fixations was conducted over this time window, with picture (with the two levels 'confusable vowel' and 'distinct vowel') as the within-participants factor. In the item analysis picture was the between-items factor. Competitors with confusable vowels were fixated significantly more often than competitors with distinct vowels ($F_1[1, 19] = 11.80$, $p < .01$; $F_2[1, 19] = 8.01$, $p < .01$). This suggests that phonetic discrimination difficulties in the non-native language lead to prolonged activation of competitors with confusable vowels during non-native spoken-word recognition. Competitors are still activated during a period (i.e., the measured 292 ms of putative overlap) when acoustic information should count against them as possible candidates.

Prior to the point that fixations could be driven by acoustic input, some variation between fixation proportions for targets, competitors, and unrelated distractors was observed (Figs. 2a and b) suggesting that the pictures, or their placement, were in some way intrinsically less attractive. Analyses in the 0–300 ms window showed significant differences in fixation proportions for participants, though not for items. For trials with confusable vowels (Fig. 2a), unrelated distractor pictures were disfavored and for trials with distinctive vowels target pictures were disfavored (Fig. 2b). To affirm that initial biases had no implications for target and competitor fixations later on, we conducted additional analyses in which trials were only included on which participants did not fixate either target or competitor at the onset of the target noun (65% of the trials). In the 300–800 ms time window, Dutch listeners again fixated competitors with confusable vowels significantly more than competitors with distinct vowels ($F_1[1, 19] = 5.02$, $p < .05$; $F_2[1, 19] = 3.63$, $p > .05$).

Note that for trials with confusable vowels (Fig. 2a), the fixation proportions to the target rise rather slowly until 500 ms after target word onset. During this time frame available acoustic information of the target word is clearly being processed, since the fixation proportions to the competitor can be seen to start rising approximately 200 ms earlier. Over a 300–500 ms time window this difference was significant by participants though not by items ($F_1[1, 19] = 5.32$, $p < .05$; $F_2[1, 9] = 1.02$, n.s.). This pattern implies that at first only *pencil* was activated when *pan-* was heard, not *panda*. Only once subsequent phonetic information mismatched the

competitor did the fixation proportions to the target word increase at a faster rate. This issue will be further considered in the General Discussion.

Experiment 1b

Two trials were removed from the analysis because no fixation on the target objects could be found even though the listeners had clicked on the correct object (0.8% of all trials). Fig. 2c presents the proportions of fixations averaged over participants for trials with confusable vowels and Fig. 2d for trials with distinct vowels. Fixation proportions for the two unrelated distractors were again averaged. Fig. 3b again contrasts the proportions of fixations for the two competitors.

The fixation proportions in Fig. 3b show that English listeners looked as often and long at the competitor in trials with vowels that are confusable for Dutch listeners as they did in trials with vowels that are distinct for Dutch listeners. As in Experiment 1a, between 0 and 300 ms there was no significant difference between fixation proportions to the two competitors ($F_1[1, 11] = 1.38$, n.s.; $F_2 < 1$). A direct comparison of the further course of fixation proportions seemed therefore justified. In the time window from 300 to 800 ms, 16.2% of the fixations were to the competitor in trials with confusable vowels and 15.3% to the competitor in trials with distinct vowels. In ANOVAs comparable to Experiment 1a, there was no significant difference in fixation proportions (F_1 and $F_2 < 1$). Thus native English listeners, who should have no difficulties discriminating between the tested vowels, indeed show no prolonged activation of competitors containing these vowels. The results of Experiment 1b thus support the proposal that prolonged activation of certain competitors by Dutch listeners stems from phonetic discrimination difficulties for Dutch listeners hearing English.[3] Together, the results of Experiment 1 strongly suggest that the inability of non-native listeners to make phonemic distinctions as rapidly and surely as native listeners leads to an increase in the effective competitor population during word recognition.

However, the tested discrimination difficulties are not symmetrical for Dutch listeners, at least in phonetic identification tasks. Although English /æ/ is often reported as /ɛ/, /ɛ/ is less likely to produce the response /æ/ (Cutler et al., forthcoming). The same is true for /aɪ/-/eɪ/: /aɪ/

---

[2] In the eye-tracking literature, ending points for lexical competition effects vary somewhat between 700 and 1000 ms (e.g., Allopenna et al., 1998; Dahan et al., 2000, 2001). We have chosen 800 ms because across all experiments it best described the ending point of competition effects, and using a fixed time window facilitated comparison of results across experiments.

[3] To check that the results of Experiment 1a were also robust for a smaller sample size, comparable to that of Experiment 1b, the results of the Dutch participants in Experiment 1a were randomly split in two groups of 10 participants each. Statistical analyses revealed comparable competition effects for both groups. The relatively small number of British English participants in Experiment 1b made analyses with exclusively trials on which participants did not fixate either target or competitor at target noun onset inappropriate.

receives an /eɪ/ response more often than vice versa. Phonetic identification tasks may, of course, not provide a reliable guide to phonetic processing for word activation and competition; it may be that Dutch listeners' identification decisions exhibit a response bias towards those categories (/ɛ/, /eɪ/) which more closely approximate available Dutch categories. Such a bias could in principle be found even though discriminability was symmetrically distributed.

Nonetheless, it may also be the case that the identification results do provide a reliable guide to discrimination in on-line listening, and that listeners are better able to resolve competition of /æ/ with /ɛ/ than of /ɛ/ with /æ/. In that case, we should also observe that *pencil* receives less competition from *panda* than *panda* received from *pencil* in Experiment 1; in fact, the confusable and distinct vowel pairs of Experiment 1 may pattern analogously with respect to competitor activation. In Experiment 2, therefore, the materials of Experiment 1 were again presented to Dutch listeners, but this time the role of target and competitor was switched for pairs with confusable vowels (e.g., *pencil* was now the target and *panda* its competitor).
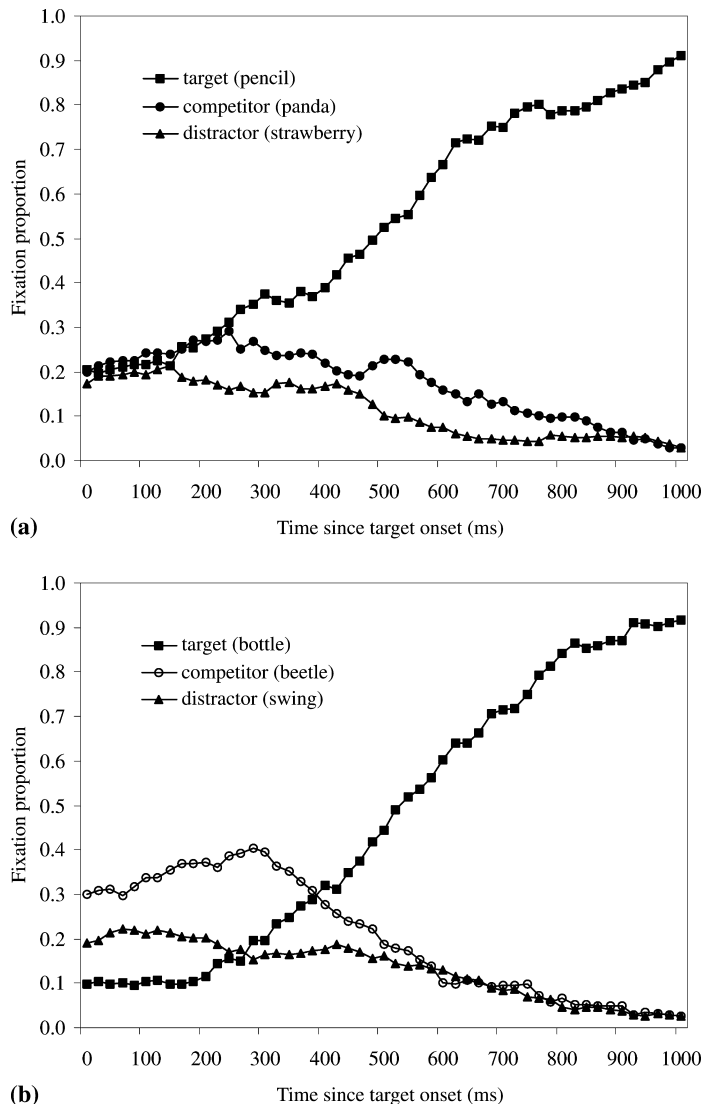

Fig. 4. Experiment 2. Fixation proportions of Dutch listeners over time for English targets, competitors, and averaged distractors (Materials as in Experiment 1 but role of target and competitor switched): (a) confusable pairs from Experiment 1 and (b) distinctive pairs from Experiment 1.

Experiment 2

Method

Participants

Twenty native speakers of Dutch, students at the University of Nijmegen, were paid to take part. They had not participated in the earlier experiment. Again, they had lived in the Netherlands all their lives, and had normal or corrected-to-normal vision and normal hearing. At a mean age of 11 they had started learning English as a foreign language at school for an average of 8 years. Participants underwent the same multiple-choice test in English as Dutch participants in Experiment 1. Their average of correct answers was 99%.

Materials

The materials were as in Experiment 1, except that the role of target and competitor was switched for pairs with confusable vowels. For example, *panda* had been a target in Experiment 1 (and therefore actively named in the experiment) and *pencil* its (unnamed) competitor. In Experiment 2, *pencil* became the target and *panda* its competitor. Pairs with distinct vowels were not switched (e.g., target *bottle* and competitor *beetle*). All unrelated distractors and filler trials were as in Experiment 1.

The spoken instructions were recorded during the same session as the material of Experiment 1, by the same male native speaker of British English. An instruction was for example: *Click on the pencil. Now put it on top of the circle.* The average duration of the preceding context (*click on the*) was 384 ms, and that of the target word was 555 ms for targets in confusable target–competitor pairs (e.g., *pencil*) and 516 ms for targets in distinct target–competitor pairs (e.g., *bottle*). In addition, the duration of the phonemic overlap between the target word and its competitor (e.g., of /pɛn/ in *pencil*) was measured. The average duration of overlap was 234 ms for confusable target–competitor pairs and 386 ms for distinct target–competitor pairs.

Procedure

The procedure was as in Experiment 1.

Results and discussion

The target–competitor pair *pedals–paddle* was excluded from the analyses due to a recording error. *Pedals* had been recorded in the singular form but the accompanying picture showed two pedals. Six more trials were removed because participants clicked on objects other than the target, and one trial because no fixation on the target object was found (1.8% of the remaining trials). Fig. 4 presents the proportions of fixations averaged over participants for trials with confusable vowels in the target and its competitor (Fig. 4a)

and for trials with distinct vowels (Fig. 4b). Fixation proportions for the two unrelated distractors were averaged. Fig. 5 shows the proportions of fixations for the two competitors from Fig. 4.

In Fig. 5, it can be seen that the fixation proportions to neither of the two competitors starts to rise at about 300 ms in the way the competitor fixations rise in Fig. 3a. In other words, Dutch listeners here did not fixate competitors of the pairs that in Experiment 1 had proved confusable more often or longer than competitors of distinct target–competitor pairs. Over the 300–800 ms time window fixation proportions were 18% to both competitors with confusable and distinct vowels. A one-factor ANOVA with picture (with the two levels 'confusable vowel' and 'distinct vowel') as the within-participants and between-items factor showed no significant difference in fixation proportions (F_1 and $F_2 < 1$). Also for a shorter time window from 300 to 500 ms no significant difference in fixation proportions was found ($F_1[1, 19] = 3.28$, $p > .05$; $F_2[1, 18] = 1.14$, n.s.).

Unlike in Experiment 1, fixation proportions to the two competitors here differed in the time window from 0 to 300 ms. There were fewer fixations on the confusable competitor (e.g., *panda*) than on the distinct competitor (e.g., *beetle*). This difference was significant by participants though not by items ($F_1[1, 19] = 19.75$, $p < .001$; $F_2[1, 18] = 1.25$, n.s.).

Again there was some variation in fixation proportions between 0 and 300 ms. Initial fixation proportions were low for the target in distinctive pairs, as they were also in Experiment 1a (Fig. 2b), and high for the competitor (Fig. 4b). We once more conducted additional analyses in which only trials were included on which participants did not fixate either target or competitor at the onset of the target noun (60% of the trials). In the 300–800 ms time window again no significant difference between fixation proportions to different competitors was found ($F_1 < 1$; $F_2[1, 18] = 1.29$, n.s.).[4]

In contrast to Experiment 1 (Fig. 2a), no delay was found here in the rise of target activation compared to competitor activation for trials with confusable vowels (Fig. 4a). Three hundred millisecond after target word onset, fixation proportions to the target start to rise and simultaneously, fixation proportions to the competitor

---

[4] In all our experiments, we had the impression that as with reading text in lines, participants were more likely to fixate objects in the upper two cells first. Even though placement of pictures was randomized, occasionally certain picture types were placed more often in the lower or upper cells than others. Initial fixation biases may be partly due to such asymmetries. At least for trials with distinct competitors in Experiments 1 and 2, asymmetries in picture placement seem to match initial preferences for picture types. We emphasize however that wherever listeners happened to be looking when the speech input started, processing of that input always produced a clear rise in fixations to the target and/or competitor.
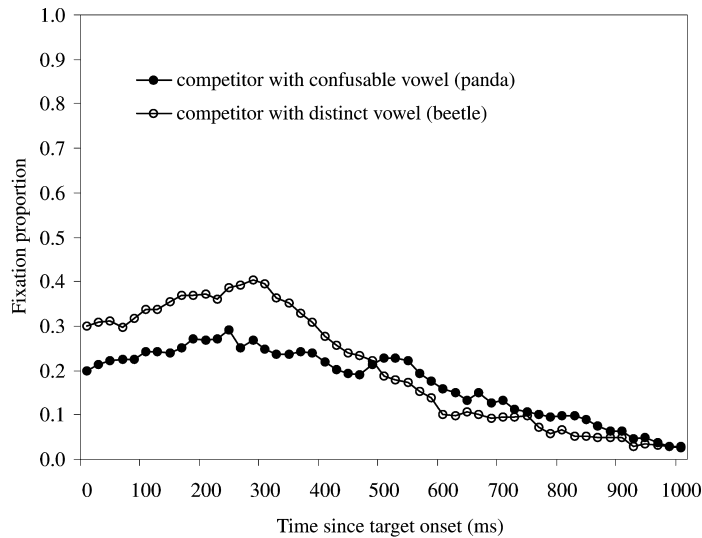
Fig. 5. Experiment 2. Fixation proportions of Dutch listeners over time for English competitors when target vowel was confusable (Fig. 4a) and when target vowel was distinctive (Fig. 4b).

start to fall. Thus, _pencil_ was immediately activated when _pen-_ was heard. The asymmetry of Dutch listeners' discrimination difficulties in English seems to control lexical activation. Regardless of whether Dutch listeners encounter English /ɛ/ or /æ/, only words with the vowel closest to a Dutch category, i.e., words with /ɛ/, are initially activated. This issue will be considered in detail in the General discussion.

The results of this experiment show that phonetic confusability (in the case of the two vowel contrasts we have tested) can have asymmetric effects on word activation and competition. Dutch listeners are more likely to identify English /æ/ as English /ɛ/ than vice versa, and they are more likely to identify English /eɪ/ as English /aɪ/ than vice versa; directly in consequence, they are more likely to activate English words with /ɛ/ when they hear English words with /æ/ than vice versa, and they are more likely to activate English words with /aɪ/ when they hear English words with /eɪ/ than vice versa. This finding has implications for theories of second-language word recognition.

Models of phonetic perception in a second language (e.g., Best, 1995) predict discriminability of phoneme categories by non-native listeners by reference to the relationship of the phoneme repertoires of the first and second language in question. Such models predict listener performance in phonetic discrimination tasks with considerable success. However, the effects of poor discriminability on word recognition may not be directly predictable from these kind of tasks. Our finding that patterns of phoneme confusion causing competitor activation can be asymmetric, so that some kinds of competitors cause more interference in listening than

others, will be considered within this framework in the General Discussion. First, however, we attempt to obtain more relevant evidence by considering the role of the native vocabulary (and by implication the phonemic categories of the native language).

Note that despite the asymmetry of competitor activation it is clear from our findings that second-language listeners can experience unwanted competition as a result of phoneme processing difficulty. Native listeners were less distracted by a picture of pencil when looking for a panda than non-native listeners were; the latter thus experienced more competition than was the case for the native-language listeners.

As described in the introduction, another way in which non-native listeners may suffer from more competition than occurs for native listeners is that native-language competitors may be spuriously activated as well. Spivey and Marian (1999) found that native speakers of Russian experienced competition from competitors with Russian names similar to the English target names. This further complicates the situation for non-native listeners; more than one factor may thus prompt unwanted increases in lexical competition. In our next experiment, we explore the native vocabulary effect for our Dutch listening population, in conjunction with the phoneme discriminability effect which we have established in the experiments so far.

Note that our population differs in a number of ways from that tested by Spivey and Marian. First, the pair of languages in question is Dutch and English rather than Russian and English. Dutch is much more closely related to English than Russian is. This may

affect the likelihood of simultaneous activation of the two vocabularies (arguably either increasing it due to the existence of many cognates, or decreasing it due to past experience of many "false friends"). Further, as we have already considered, the phonetic repertoires of Dutch and English exhibit mismatches, and this has consequences for activation. Spivey and Marian do not discuss the relationship between the phonetic repertoires of Russian and English. It may be that greater phonetic mismatch leads to lesser likelihood of concurrent multi-vocabulary availability. Finally, and most importantly, the listeners tested by Spivey and Marian were not resident in their native country but in the country of their second language. The dominant requirement for daily communication was therefore use of English rather than Russian. The dominant use of a second-language vocabulary may encourage a more symmetrical and concurrent availability of the native and second-language vocabulary. In contrast, asymmetric use of the first and second language may make it easier to keep vocabularies apart. This latter case applies to our listener population, who were resident in their native country, and, though adept in the use of English, definitely used much more Dutch than English every day. For all these reasons it is far from obvious that the native-language interference observed by Spivey and Marian will be observed with Dutch listeners listening to English.

Of course, it may also be the case that native-language interference is always present and over-riding, and further that far greater use of the native than of the non-native vocabulary exacerbates it; in this case we should observe even more marked effects than Spivey and Marian found. Experiment 4 will provide an answer to these questions. In addition, however, Experiment 4 allows us to extend the question to the role of phoneme confusability in activation of spurious native competitors. If native phoneme categories capture non-native phonemes, do nominally mismatching as well as better-matching word overlaps between native and non-native words result in competition? That is, will we find not only that English *kitten* might suffer unwanted competition from Dutch words such as *kist*, 'chest,' containing essentially the same /ɪ/ vowel, but that English *carrot* /kærət/ might also suffer competition from Dutch words beginning with /kɛ/ such as *kerk*, 'church'?

In Experiment 3, then, native Dutch listeners were presented with spoken English words while the visual display from which they had to select a target included a distractor item of which the Dutch (but not the English) name made it a potential competitor. If the native vocabulary interferes in listening to the second language for the present listeners too, despite the potentially modulating factors listed above, then we will observe that these distractors cause more interference than other

distractors with names unrelated to that of the target in either language. Further, some of the Dutch names involved a vowel matching closely to an English vowel (e.g., /ɪ/, /i/) while others involved a confusable contrast of the kind examined in Experiments 1 and 2 (e.g., /æ/, /aɪ/). If the phonetic confusions we observed in Experiment 1 carry through to activation of native competitors exemplifying the relevant native phonemic categories, then we will observe that both sets of competitors cause interference. However, if second-language categories are confusable but not thereby necessarily perceived as the native category, we will observe that the better-matching competitors (*kist* given *kitten*) cause greater interference than the mismatching candidates (*kerk* given *carrot*). Like Experiment 1, Experiment 3 also included a control group of listeners for our Dutch population. It is possible that the pictures with Dutch names similar to the English target names could be intrinsically more attractive. Such an effect should, however, appear with any participant group, including English listeners who do not know any Dutch. However, since there is no reason to expect that the English word recognition by such a group would be subject to competition from Dutch words, in the absence of such an effect we should observe a different pattern for the native listeners: Distractors with potentially competing Dutch names should offer no more interference than that provided by other distractors.

Experiment 3

Method

Participants

Twenty students from the University of Nijmegen took part in Experiment 3a. As for Experiments 1 and 2, these were native speakers of Dutch who had lived in the Netherlands all their lives, and had normal or corrected-to-normal vision and normal hearing. They had received on average 7.5 years of training in English as a foreign language in secondary education, beginning at a mean age of 11. None had participated in Experiments 1 or 2. Participants underwent the same English multiple-choice test which Dutch participants in Experiments 1 and 2 took after the experiment. The average score was 97% correct. Ten native speakers of American English participated in Experiment 3b. At the time of testing none of the participants had been in the Netherlands for more than 5 days. They had no knowledge of Dutch. All participants received a small payment for taking part.

Materials

The target words consisted of 20 English nouns referring to picturable objects (e.g., *kitten*). Each English

target word was paired with a Dutch competitor. The onset of the competitor in Dutch overlapped phonemically with the onset of the target word in English (e.g., English target word *kitten*, Dutch competitor *kist*, 'chest'), but there was no other onset overlap within or between items. Thus, the name of the target item in English overlapped neither with the name of that item in Dutch (e.g., *kitten* is *poesje* /pusjə/ in Dutch) nor with the English name of the Dutch competitor (e.g., *kist* in English is *chest*). As in Experiment 1, two phonologically unrelated distractors were added for each target word (e.g., *flower* and *swing*). Neither the English nor the Dutch names of the unrelated distractors (e.g., *bloem* /blum/, 'flower' /flaʊər/ and *schommel* /sxɔməl/, 'swing' /swɪŋ/) overlapped with the English target word. Seven of the 20 target words had a potential Dutch competitor with a closely approximating vowel to that in the English target (e.g., target word *kitten* /kɪtn/ and Dutch competitor *kist* /kɪst/, 'chest'); the remaining 13 contained a vowel which the English vowel matched less well, but which was confusable, as shown in Experiment 1 (e.g., target word *carrot* /kærət/, Dutch competitor *kerk* /kɛrk/, 'church'). The 20 English target words, their Dutch competitors, and unrelated distractors are listed in Appendix B. Twenty further filler trials were constructed, under the same constraints as in Experiment 1.

Pictures comparable to those used in Experiment 1 were selected. When 10 native speakers of Dutch were asked to name the target and competitor pictures in Dutch and English, the agreement between participants' responses and the intended names was 91% in Dutch and 85% in English. An additional 10 native speakers of Dutch were asked to rate the goodness of the pictures as pictures of the intended object on a scale from 0 to 7. They rated the goodness of the pictures with a mean of 5.8. Again, some small suggestions for improvement of the pictures were implemented.

In addition, lexical frequencies of targets and of competitors were counted using the *CELEX* database (Baayen et al., 1993). To compare the conceptual frequency of the target and competitor items we computed lemma frequencies in Dutch of the targets (e.g., *bureau*, 'desk'; mean 39.71 per million) and the competitors (e.g., *deksel*; mean 37.07 per million). To compare the frequency of the form actually heard against its putative form competitor we ascertained English word form frequencies of the targets (e.g., *desk*; 42.01 per million) and Dutch word form frequencies of the competitors (e.g., *deksel*; 24.55 per million). For 13 of the 20 pairs, the name of the English target had a higher frequency than the name of the Dutch competitor in the word form count. Statistical analyses revealed no significant difference in any comparison (all $Fs < 1$).

The spoken instructions were recorded, by a male native speaker of American English, in the same manner as for Experiment 1. An instruction was for example: *Click on the kitten. Now put it on top of the diamond.* The average duration of the preceding context (*click on the*) was 451 ms, and of the target word (e.g., *kitten*) 575 ms. In addition, the duration of the phonemic overlap between the English target word and its Dutch competitor (e.g., the duration of /kɪ/ in *kitten*) was measured for all 20 items. The average duration of overlap was 270 ms. When the seven items with closely approximating vowels were excluded, phonemic overlap was 266 ms.

Procedure

The procedure was as in Experiment 1. Participants were not made aware of potential cross-language competition in the experiment.

Results and discussion

Experiment 3a

On six trials participants erroneously clicked on an object other than the target object (1.5% of all trials). These trials were excluded from the analyses. Fig. 6a presents the proportions of fixations averaged over participants to the target, the competitor, and the average for the two unrelated distractors after target onset.

Fig. 6a shows that the probability of fixating the Dutch competitor began to diverge from the probability of fixating the unrelated distractors about 300 ms after target word onset. The probability of fixating the Dutch competitor remained greater than that of the unrelated distractors until approximately 800 ms after target word onset. To compare the proportions of fixations to the competitor and to the average for the two unrelated distractors, a time window extending from 300 to 800 ms after target onset was defined. Over this window the proportion of fixations was 23.6% for the Dutch competitor and 16.5% for the average of the unrelated distractors. A one-factor ANOVA on the mean proportion of fixations over the 300–800 ms time window, with picture (with the two levels 'competitor' and 'unrelated distractors') as the within-participants factor, showed that the competitor was fixated significantly more often than the unrelated distractors ($F_1[1, 19] = 15.62$, $p < .001$; $F_2[1, 19] = 5.54$, $p < .05$). This suggests that during the presentation of the English target words the Dutch competitors were activated.

An additional one-factor ANOVA over the same time window (300–800 ms) was performed excluding the seven items with closely approximating vowels (e.g., *kitten*). The mean proportion of fixations was 22.9% for the Dutch competitor and 18.2% for the average of the unrelated distractors. The difference was significant by participants ($F_1[1, 19] = 5.00$, $p < .05$), but did not quite reach significance by items, probably due to the relatively low number of items ($F_2[1, 12] = 3.63$, $p > .07$). Thus, evidence for the
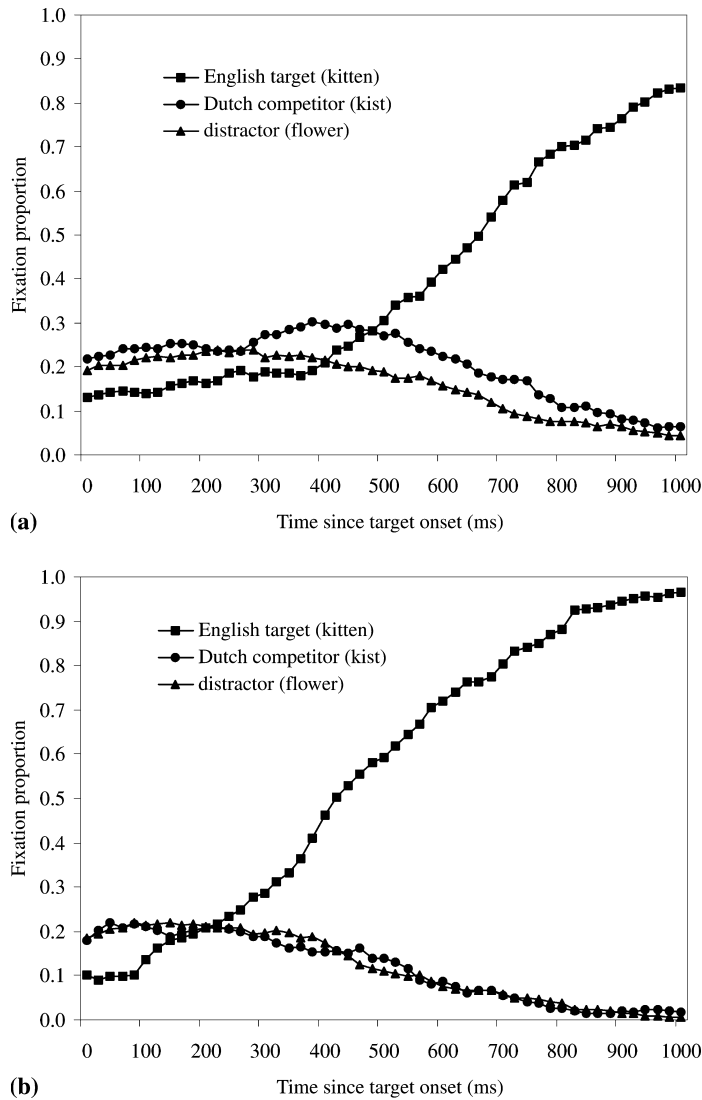
Fig. 6. Experiment 3. Fixation proportions over time for English targets, Dutch competitors, and averaged distractors: (a) Dutch listeners and (b) American English listeners.

activation of Dutch competitors appeared even when the vowel of the English target words matched less well with the vowel of the Dutch competitors.

To examine differences in fixations to pictures before any acoustic information from the target word could influence eye movements, an ANOVA was conducted on the fixations to the target, the competitor and the average of the unrelated distractors over a time window extending from 0 to 300 ms. The fixation proportions over the first 300 ms after target onset differed significantly only by participants ($F_1[2, 38] = 8.45$, $p < .001$; $F_2[2, 38] = 1.33$, n.s.). Newman–Keuls tests indicated that the proportion of fixations to the target was lower than that to both the competitor and the unrelated distractor, but the proportion of fixations to the competitor did not differ signifi-

cantly from that to the unrelated distractor. This suggests that the difference found between fixations to the competitor and to the unrelated distractor in the 300–800 ms time window cannot be attributed to a general bias toward the picture of the competitor.

Experiment 3b

Five trials were removed from the analysis (2.5% of all trials) because participants either clicked on an object other than the target object (2 trials) or no fixation on the target object was found (3 trials). Fig. 6b presents the proportions of fixations averaged over participants to the target, the competitor, and the average for the two unrelated distractors after target onset.
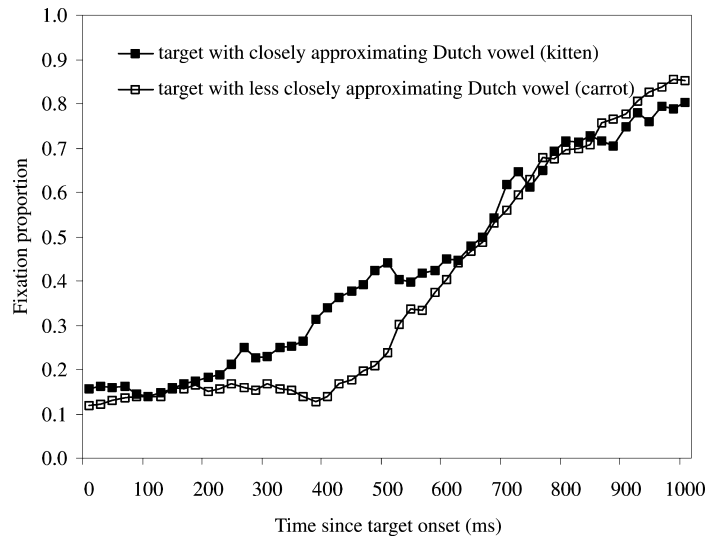
Fig. 7. Experiment 3. Fixation proportions of Dutch listeners over time for English targets, separately for targets with vowels closely approximating a Dutch vowel, and targets with vowels less closely approximating a Dutch vowel.

As can be seen in Fig. 6b, the fixation proportions of the 10 American participants were equivalent for the Dutch competitor and the averaged unrelated distractors. Unlike the Dutch participants, American participants did not look more often at the Dutch competitor than at the unrelated distractors.[5] As for Experiment 3a, a time window of 300–800 ms was compared; in this window, the proportion of fixations was 10.7% for the Dutch competitor and 10.9% for the average of the unrelated distractors. The one-factor ANOVA with picture as the within-participants factor showed that there was indeed no significant effect (F_1 and $F_2 < 1$). This suggests that during the presentation of the English target words the Dutch competitors were not activated, which in turn suggests that the results for the Dutch listeners reflect competition from the native vocabulary.

Note one further difference between the non-native and native results in this experiment: Fig. 6a shows an advantage for the competitor over the target until approximately 500 ms after target word onset, while no such effect appears in Fig. 6b. Over a 300–500 ms time window this difference for the non-native listeners was significant by participants though not by items ($F_1[1, 19] = 9.69$, $p < .01$; $F_2 < 1$). Between 300 and 400 ms the fixation proportion of the competitor rises, indicating that the incoming acoustic information is being processed at that time. The fixation proportion of the target, however, does not start to rise until about

400 ms, at which point the fixation proportion of the competitor starts to fall. This suggests that the target is only considered as a candidate at 400 ms and beyond. The relatively low proportion of fixations to the target could partly be due to the general bias against the target pictures, that also appears prior to target word onset. The difference between fixations to the target and the competitor might also be a subjective frequency effect across languages. Although no significant difference obtained between the lexical frequency of targets and competitors, Dutch participants presumably have heard the Dutch word *kist* more often in their lives than the English word *kitten*. In this sense, the Dutch competitors had a higher frequency than the English targets. However, recall also that in Experiment 1a a similar advantage for the competitor over the target was observed for target–competitor pairs with confusable vowels (Fig. 2a). English target words with the vowel /æ/ initially activated English words with /ɛ/ more often than English words with /æ/. Similarly, in Experiment 3a, English target words with vowels that match less closely to a Dutch category (e.g., *carrot*) may have been initially disregarded in lexical activation, whereas English targets with vowels that closely match (e.g., *kitten*) might have been activated immediately.

If this hypothesis is correct, then the delay of target activation here should be due to the former set of words, not the latter. Fig. 7 plots fixation proportions to the targets separately for these two subsets of the materials. Fixation proportions in Experiment 3a were, as we noted, initially lower for targets than for distractors or competitors. Fig. 7 shows that there is, however, no difference in the fixation proportions for

---

[5] To check that the results of Experiment 3a were also robust for a smaller sample size, results were randomly split in two groups of 10 participants each. Statistical analyses revealed comparable competition effects for both groups.

the two types of targets, respectively, in the initial 0–300 ms range. Nevertheless, the two types of target differ in when they start to rise from this initial equivalent state. For targets with closely approximating Dutch vowels, fixation proportions start to increase around 300 ms. This simply does not happen for targets with less closely approximating Dutch vowels; their rise starts about 100 ms later. This is clearly not due to initial differences, but to the nature of the vowel. The difference in the 300–600 ms window is significant across participants ($F_1[1, 19] = 15.29$, $p < .001$; $F_2[1, 18] = 3.35$, $p > .05$). Further implications of this will be considered in the General Discussion.

The results of Experiment 3 strongly suggest that non-native listeners experience spurious competition from native candidates during word recognition in the non-native language, even though the native language is irrelevant for the task. This effect appears with listeners whose dominant use of the native vocabulary might have encouraged clear vocabulary separation, just as earlier work showed it to hold for listeners who were required to maintain dominant use of the non-native vocabulary. Moreover, the degree of match between native and non-native phonemic categories did not strongly modulate this competition effect: Competitors based on a less close match (*kerk* given *carrot*) provided interference just as competitors based on a closer match (*kist* given *kitten*) did. This suggests that the second-language input capture by native phonemic categories translates directly into native-language lexical activation.

English spoken-word recognition by native English listeners could not be subject to competition from candidate words of Dutch, a language they were unfamiliar with. Thus the difference in the pattern of results across Experiments 3a and 3b can only be attributed to a cross-language difference in processing. The results for the Dutch listeners do not reflect some artifact of the material: they reflect the native language of the participants and its propensity to offer competition even with spoken input for which it is clearly irrelevant.

In our final experiment, we further tested the implications of the differences between our listener population and that tested by Spivey and Marian (1999) and Marian and Spivey (1999, 2003). As described in the introduction, they were able to observe interference from the second language in listening to their native language, in their population of listeners resident in the second-language country and required to maintain dominant use of the second-language vocabulary. Our listeners are resident in their native country, and use their second language much less than their native language. We therefore tested whether they too would experience competition from the second language when listening to the native language—that is, whether English words like *kitten*, which Experiment 3 clearly showed to

be known to these participants, would cause competition for detection of the Dutch target *kist*. For Experiment 4, the materials of Experiment 3 were translated into Dutch and presented to native Dutch listeners. As in the earlier experiments, these listeners were highly proficient in English, but here they were not required to exercise this proficiency.

Experiment 4

Method

Participants

Twenty native Dutch speakers took part, in return for a small payment. They had not participated in any of the earlier experiments. During their secondary education they had received on average 6.95 years of training in English as a foreign language, beginning at a mean age of 11.

Materials

The Experiment 3 materials were translated into Dutch. Competitors of Experiment 3 became targets and were now named in the experiment (e.g., *kist*, 'chest,' was now a target). In turn, targets became competitors and were thus not named (e.g., *kitten* was now a competitor). The spoken instructions were recorded by a male native speaker of Standard Dutch. An instruction was for example: *Klik op een kist. Plaats hem nu op de ruit.* 'Click on a chest. Now put it on top of the diamond.' Note that the Dutch indefinite article *een* preceded the target noun. Definite articles would have been marked for gender in Dutch; as Dahan et al. (2000) have shown, a gender-marked article can suppress early activation of gender-inconsistent competitors. The average duration of the preceding context (*klik op een*) was 478 ms, and that of the target word (e.g., *kist*) was 522 ms. The average duration of overlap between the Dutch target word and its English competitor (e.g., the duration of /kɪ/ in *kist*) was 238 ms.

Procedure

The procedure was as in Experiment 3, except that the participants were told that they would hear Dutch, and instructions were given in Dutch.

Results and discussion

Five trials were removed from the analysis (1.25% of all trials). In one trial, a participant had clicked on an object other than the target object, in the other four trials no fixation on the target object could be found. Fig. 8 presents the proportions of fixations averaged over participants to the target, the competitor, and the average of the two unrelated distractors after target onset.
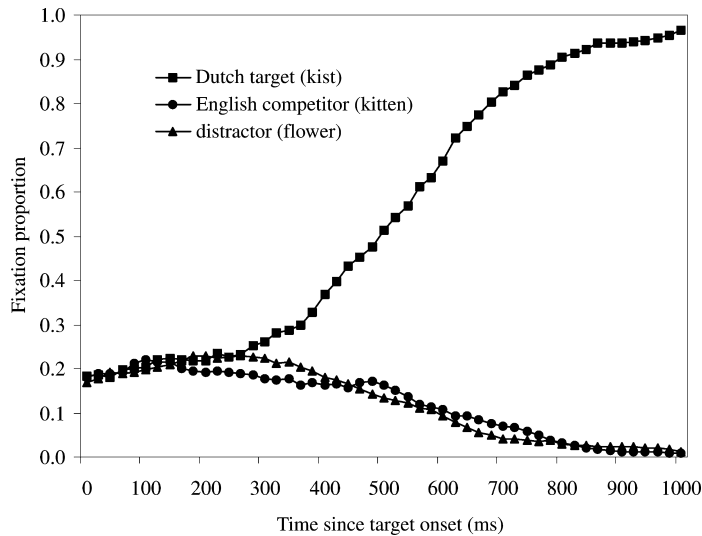
Fig. 8. Experiment 4. Fixation proportions of Dutch listeners over time for Dutch targets, English competitors, and averaged distractors.

As is immediately apparent from Fig. 8, the fixation proportions to the English competitor and to the average of the unrelated distractors do not differ. Over the 300–800 ms time window the proportion of fixations was 12.2% for the English competitor and 11.8% for the average of the unrelated distractors. That is, these Dutch participants did not look more often at the English competitor than at the unrelated distractors. A one-factor ANOVA on the mean proportion of fixations with picture (with the two levels 'competitor' and 'unrelated distractors') as the within-participants factor, confirms the lack of a difference in viewing times (F_1 and $F_2 < 1$). The pattern of results suggests that Dutch listeners listening to Dutch target words did not activate candidate words in their second language, English.

If English competitors were activated in Experiment 4, it would be likely that the competitor activation is at most very brief for items in which the vowels of the target and competitor match less well (e.g., target *kerk*, competitor *carrot*). In Experiment 2, Dutch listeners fixated *panda* after *pen-* for no longer than they fixated distinct competitors. Thus in Experiment 4, English competitors with /æ/ are unlikely to be activated by Dutch targets with /ɛ/. However, target–competitor pairs with closely approximating vowels (e.g., target *kist*, competitor *kitten*) might still possibly show activation of English competitors. For the target–competitor pairs with closely approximating vowels, 14.2% of the fixations were to the competitor and 14.0% to the average of the unrelated distractors in the 300–800 ms time window (F_1 and $F_2 < 1$). Clearly, no activation of English competitors was evident.

Note that a subjective frequency effect across languages might again have played a role. As we pointed out above, Dutch participants presumably have heard the relevant Dutch words more often in their lives than the relevant English words, so that the English competitors had a much lower frequency than the Dutch targets. This may be one reason why our results in Experiment 4 contrast with the findings of Spivey and Marian (1999) and Marian and Spivey (1999, 2003) who found not only lexical activation of the native language interfering with processing of the non-native language, but also interference of the non-native language during processing of the native language. We discuss the further implications of this finding below.

General discussion

Our four experiments motivate a gloomy view of spoken-word recognition in a second language: The available vocabulary in that language may be smaller, but the extent of lexical competition is greater. The more competition in word recognition, the slower recognition is accomplished—therefore, non-native listeners may be doomed to recognize spoken language less rapidly than native listeners.

Our results revealed two potential sources of added competition. First, non-native listeners' phonetic discrimination difficulties cause inappropriate competitor activation. Dutch participants in Experiment 1 fixated distractor pictures more and longer when the English names of the target and distractor picture contained vowels that for Dutch listeners are confusable than when they contained distinct vowels. Native English listeners presented with the same materials showed no such difference in viewing times. Second, the non-native listen-

ers' native vocabulary adds competition. Dutch listeners in Experiment 3 fixated distractor pictures of which the Dutch name resembled the English name of the target picture. Native English listeners with no knowledge of Dutch showed no activation of these Dutch pseudo-competitors.

Both of these effects are unidirectional. When the stimuli of Experiment 1 were again presented to Dutch listeners, but with the role of target and competitor switched for confusable-vowel pairs, Dutch listeners (in Experiment 2) showed no prolonged competitor activation for these pairs. The confusability of the vowel contrasts that we tested is itself asymmetric; this asymmetry thus carries through to word recognition, in that *pan-*, for instance, will activate *pencil*, but *pen-* will not activate *panda*. Finally, Experiment 4 showed no interfering activation of phonologically related English competitor words for Dutch listeners listening to their native language. This result contrasts with findings by Spivey and Marian (1999); (Marian & Spivey, 1999, 2003) who report for some of their experiments activation of non-native competitors while listening to the native language. In our results cross-vocabulary competition appeared specific to non-native listening.

Most, perhaps all second-language learners know from personal experience that recognizing non-native spoken utterances can present difficulty; the contrast with the familiar ease of native listening is striking. Of course, non-native listeners can be confronted by words and expressions that they have never before encountered in their second language, and this problem certainly underlies a good part of the difficulty confronting such listeners. But in theory a small vocabulary might provide known lexical items with less competition than they would experience from neighbors in a larger vocabulary. However, we have shown that any such benefit is certainly compensated for, and perhaps even overwhelmed, by the additional competition generated from two sources: on the one hand, phonetic discrimination difficulties carrying through to lexical activation, and on the other hand the availability, even in monolingual second-language situations, of the members of the native vocabulary. Together these factors provoke significant spurious competition for the non-native listener, such that the degree of competition may well exceed even that arising in the most well-stocked of native vocabularies.

The extent of the added competition effect can be appreciated by comparing the rate of target activation across experiments. Fig. 9 plots the target activation functions for the confusable pairs of Experiment 1a (non-native listeners) versus 1b (native listeners), and Fig. 10 shows the same comparison for Experiments 3a (non-native listeners) versus 3b (native listeners). In each case it can be seen that although the non-native listeners do eventually look at the target to an extent not significantly different from that of the native listeners, the rate at which they achieve this level of performance is much slower. Statistical analyses across the 300–1400 ms time frame showed significant differences both in Experiment 1 (Fig. 9; $F_1[1, 31] = 14.28, p < .001; F_2[1, 9] = 6.09, p < .05$) and in Experiment 3 (Fig. 10; $F_1[1, 29] = 59.74, p < .001; F_2[1, 19] = 66.62, p < .001$). Both phonetic discrimination difficulty and native vocabulary interference thus led to significant slowing in non-native word recognition.

Each of these two factors in its own right can potentially initiate large competition problems. Phonetic
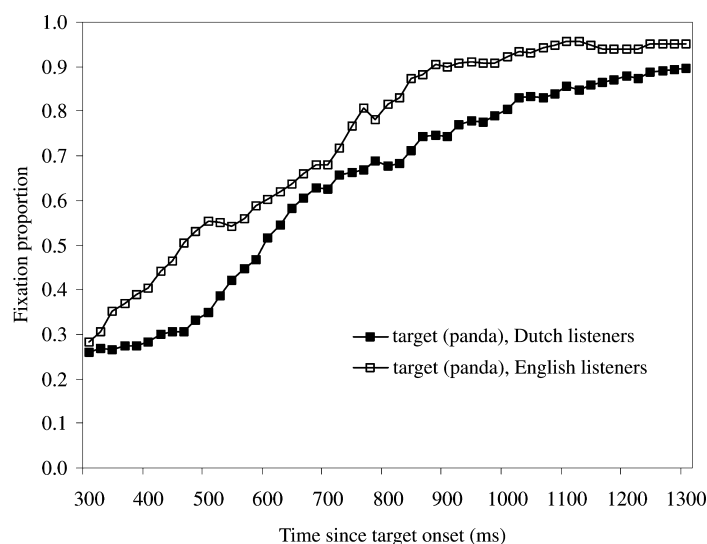

Fig. 9. Experiment 1. Fixation proportions over time for English targets with names containing vowels confusable for Dutch listeners, for non-native (Dutch) versus native listeners.
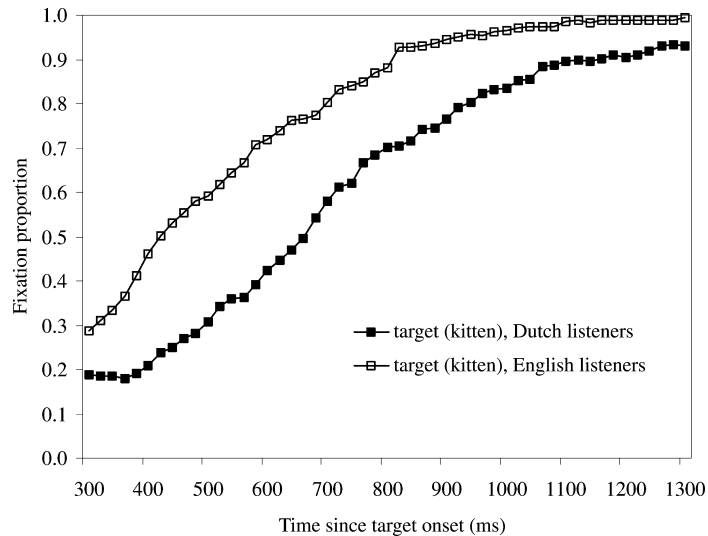
Fig. 10. Experiments 3. Fixation proportions over time for English targets for non-native (Dutch) versus native listeners.

discrimination difficulty may in principle affect the whole of a listener's second-language vocabulary, and even if the discrimination problems are confined to a single phonemic contrast, the damaging effect on word recognition could still be far-reaching. Although vocabularies are very large—tens to hundreds of thousands of words—they are made up of a comparatively tiny stock of phonetic components (around 30 phonemes in the average repertoire; Maddieson, 1984), so that it is likely that any individual phonetic contrast between two phonemes will involve thousands of words. Minimal pairs which a non-native listener cannot discriminate—such as *cattle* and *kettle* for a Dutch listener to English—may be the worst case; but as our results have shown, it is not only with such minimal pairs that word-recognition difficulty arises from the phonetic discrimination difficulty. Partial matches can lead to temporary competitor activation—of *kettle* when *catalog* is heard, of *belly* when *balance* is heard, of *pencil* when *panda* is heard, and so on through the vocabulary. These competitors may be activated only in the short term, but as previous research has shown (e.g., Norris et al., 1995), from unwanted competitor activation there ensues measurable delay in word recognition. Our results, as summarized in Figs. 9 and 10, confirm this inhibitory effect of added competition in word recognition.

The extent to which a second-language listener's phonetic discrimination problems will spread across the vocabulary depends, of course, on the exact relation between the two phonetic repertoires in question. There has been a great deal of research (see, e.g., the papers in Strange, 1995) on which types of inter-repertoire mismatches lead to serious discrimination problems, and which are relatively harmless. The most explicit

model in this area is that of Best (1995), which distinguishes a number of different possible mappings. For instance, when two categories of the second language map equally well to a single category in the listener's native language, discrimination is very difficult indeed. When one phonemic category of the second language provides a good fit to a native category while another second-language phoneme provides a poor fit to the same category, however, discrimination, though difficult, should be reasonably successful. These are the two cases in Best's classification that are most obviously relevant to the vowel contrasts manipulated in Experiments 1 and 2. Note that Best's classification has been shown to account very well for the behavior of non-native listeners in phoneme discrimination tasks, including Japanese perception of English consonants (Best & Strange, 1992), English perception of Zulu consonants (Best, McRoberts, & Sithole, 1988; Best, McRoberts, & Goodell, 2001), English perception of German vowels (Polka, 1995) and Japanese ratings of the goodness of fit of English consonants to Japanese categories (Guion, Flege, Akahane-Yamada, & Pruitt, 2000).

Despite this extensive research attention to the way phonemic repertoires may mismatch, however, there has been little work explicitly addressing the consequences of phoneme discrimination difficulties for word recognition. Pallier, Colomé, and Sebastián-Gallés (2001) demonstrated that inability to discriminate Catalan vowel contrasts caused spurious activation for Spanish-dominant Spanish–Catalan bilinguals, and Broersma (2002) showed that Dutch listeners to English accepted English nonwords such as *frash* as real words. An important outcome of the present study is the finding that discrimination difficulty, and its effect on competitor activation, may be

asymmetric. The vowel /æ/ is readily identified by Dutch listeners as the English vowel /ɛ/, but the reverse is to a lesser extent the case. When [pæ] is heard, words beginning with [pɛ] are accordingly activated, but the reverse is far less often the case. The same asymmetry appears with the diphthongs /aɪ/ and /eɪ/. Clearly, failure to discriminate a non-native contrast in phonetic processing does not rule out a role for that contrast in processing. If /æ/ and /ɛ/ were at all levels treated as instances of the same unitary category, then input of either vowel would activate words containing either vowel; but this was not the case. Instead, our results suggest that a contrast may be maintained, but effectively only one member of the contrast may play an active role. Words containing /æ/ and /ɛ/, respectively, are indeed distinguished at the level of lexical representation; in this sense Dutch listeners maintain this English contrast. But phonetic processing does not produce alternatives to map to those distinct representations; input of /æ/ and input of /ɛ/ activate the same phonetic category.

The active category for our Dutch listeners appears to be English /ɛ/. Although this category is distinct from English /æ/ at the lexical level, input of either is categorized as English /ɛ/. We suggest that this is because the native Dutch category is /ɛ/ (although the Dutch vowel with this label is not identical to the English vowel with this label). Likewise, for the /aɪ/-/eɪ/ contrast the active category is English /eɪ/, again by analogy to Dutch /eɪ/ (also not identical to the English diphthong with the same phonemic label). Non-native listeners exhibit a response bias that has its source in what response alternatives the native phonology makes available, and this bias further controls phonetic identification and, in turn, lexical activation. It is as if the phoneme category of the second language that is perceived as nearest to the native category captures all identification responses, while the second-language phoneme that is perceived as further from any native category is simply ignored. This has the extraordinary effect that words of the second language which contain that disregarded phoneme may also be, at least temporarily, disregarded in activation. As we pointed out, this is not the situation which would arise if the second-language contrast were not distinguished at all, since in such a case, input containing either phoneme would activate words containing both. Instead, the contrast is apparently represented at the lexical level. The problem then occurs at the input level: whichever vowel is actually heard, only words containing /ɛ/ are matched. No vocalic input ever matches the words containing /æ/.

We suggest that this is in fact the reason for the slow rise of looks to the designated target in Experiment 1 (see Fig. 2a). The vowel in these words seems not to have contributed to activation of the target word; substantial activation occurred, instead, only once subsequent phonetic information mismatched the competitor (for instance, the [d] of *pand*- mismatched *pencil*). Moreover, we suggest that this factor is also at work in the target

fixations in Experiment 3. Fixations to the targets with vowels approximating a Dutch alternative began to rise earlier than those to the targets containing less closely matching vowels. Thus although the purpose of Experiment 3 was to examine the activation of spurious competitors, the availability of two subsets within the materials allowed the target activations too to prove separately informative, in connection with the issues addressed in Experiments 1 and 2. Apparently, the vowel information in a word like *bike* or *carrot* was simply not effective in activating the target word in Experiment 3; instead, the input vowels in question merely activated words containing the contrasting category, and these included the native competitors (*bijl, kerk*). Activation of the second-language targets began in earnest only once further acoustic information mismatched the competitors but supported the targets (e.g., the /k/ of *bike* which mismatched *bijl*).

The capture of second-language phoneme identification by native repertoire categories provides a partial explanation of the effect of the second factor exacerbating non-native competition, namely the availability of competitors which are not words of the second language at all, but of the first. Clearly, this factor too can exercise a very substantial effect, since the entire native vocabulary can thereby participate in the competition process (and it is quite likely to be very much larger than the listener's second-language vocabulary). As Experiment 3 demonstrated, partial mismatches of the Experiment 1 type are no hindrance to activation of the native vocabulary. Superficially this may appear paradoxical, because listeners can very quickly determine the source language given a fragmentary input— even half a phoneme may suffice (Grosjean, 1988; Li, 1996). Moreover, eye-tracking studies have shown modulation of activation by the goodness of a phonemic token (McMurray, Tanenhaus, Aslin, & Spivey, 2003), and this sensitivity extends to language-specific cues, such that an English-like consonant inserted into a Spanish input causes activation of English competitor words for Spanish–English bilingual listeners (Ju & Luce, in press). Nevertheless, this discernment seemingly does not enable listeners to exclude the irrelevant vocabulary. However, the paradox is reduced if in the speech recognition process a non-native phoneme is mapped to the nearest native equivalent; the non-native phoneme may contribute to lexical activation of native words as effectively as if it were indeed the native exemplar.

Our results regarding activation of words from two vocabularies both extend and qualify the demonstrations of irrelevant lexical activation by Spivey and Marian (1999); (Marian & Spivey, 1999, 2003, in press). The most important extension consists in the detailed temporal information available in our results. Whereas Spivey and Marian's reports concern only proportions

of trials with eye movements to distractor objects, we have shed light on the time-course of lexical activation for the distractors versus the targets. We have also extended the language base of the effect to Dutch/English. But our results also qualify theirs, in that in Experiment 4 we observed no second-language word activation when participants listened to materials in their native language. The recent study of Ju and Luce (in press), mentioned above, also challenges this aspect of Spivey and Marian's findings. The Spanish–English bilinguals tested by Ju and Luce were comparable to Spivey and Marian's participants, in that they were highly proficient bilinguals living in the country of the second language. However, when these listeners were presented with native Spanish target words (e.g., *playa*, 'beach'), they did not fixate potential English competitors (e.g., *pliers*) more often than unrelated distractors. Activation of English competitors was only observed when the word-initial VOT of the initial phoneme (here, [p]) of the Spanish target names was manipulated to resemble English VOTs. This situation, in which the input contained a potentially obtrusive phonemic mismatch, is comparable to presentation of code-switched words.

There are clear differences between Spivey and Marian's testing situation and ours. First, their listeners were resident in the second-language country, and thus needed to maintain dominant use of the second-language vocabulary, while our listeners lived in their native country and there was no question of the second language dominating the first. Second, their participants were tested in both languages, while ours had no reason to suspect that more than one language was relevant for any experiment. Thus we may conclude that inter-vocabulary interference is not unlimited: At least for listeners who use their second language less frequently than their native language, competition in native listening is not increased by the existence of a second-language lexicon.

This, however, is the only ray of hope for listeners in our results. In general, the amount of lexical competition is much greater in non-native than in native listening. Spurious second-language competitors, activated as a result of difficulties in phonetic discrimination, and spurious native competitors, activated despite irrelevance of the native vocabulary in a second-language situation, both complicate the listener's task. It is their combined effect that makes non-native listening such very hard work.

Appendix A

Experimental stimuli used in Experiment 1, with phonetic transcriptions (British English)

English target words	English competitor words	English unrelated distractors	
ballot box /bælətbɒks/	belly button /bɛlibʌtn/	cloud	girl
bible /baɪbl/	baby /beɪbi/	leaf	hat
eyelid /aɪlɪd/	alien /eɪljən/	whistle	mirror
paddle /pædl/	pedals /pɛdlz/	duck	hand
palace /pælɪs/	pelican /pɛlɪkən/	leg	tomato
panda /pændə/	pencil /pɛnsl/	strawberry	dice
pineapple /paɪnæpl/	painter /peɪntə/	glasses	donkey
racket /rækɪt/	records /rɛkɔːdz/	cup	table
rider /raɪdə/	radio /reɪdɪəʊ/	shell	birdhouse
tie /taɪ/	tail /teɪl/	pot	rug
bottle /bɒtl/	beetle /bitl/	carrot	stocking
chicken /tʃɪkɪn/	chocolate /tʃɒkələt/	moustache	car
curtain /kɜːtn/	kitten /kɪtn/	swing	flower
door /dɔː/	deer /dɪə/	hairdryer	napkin
lighter /laɪtə/	letter /lɛtə/	brush	nut
lock /lɒk/	lake /leɪk/	mushroom	arrow
parrot /pærət/	pirate /paɪrət/	cap	trashcan
spoon /spun/	spine /spaɪn/	coat hanger	mouse
t-shirt /tiʃɜːt/	toaster /təʊstə/	plane	bird
violin /vaɪəlɪn/	Volcano /vɒlkeɪnəʊ/	button	scissors

Target–competitor pairs with confusable vowels are underlined. In Experiment 2, the competitor words became the target words.

Appendix B

Experimental stimuli, Experiment 3, with phonetic transcriptions (American English)

English target words and Dutch translation	Dutch competitor words and English translation	Unrelated distractors in Dutch and English	
desk /dɛsk/	deksel /dɛksəl/	bloem	schommel
'bureau'	'lid'	flower	swing
kitten /kɪtn/	kist /kɪst/	moer	borstel
'poesje'	'chest'	nut	brush
lake /leɪk/	lepel /lepəl/	muts	afvalemmer
'meer'	'spoon'	cap	trashcan
leaf /lif/	libel /libɛl/	hand	knoop
'blad'	'dragonfly'	hand	button
meat /mit/	mier /mir/	tafel	beker
'vlees'	'ant'	table	cup
seatbelt /sitbɛlt/	citroen /sitrun/	kleed	pot
'veiligheidsriem'	'lemon'	rug	pot
spring /sprɪŋ/	sprinkhaan /sprɪŋkhan/	tomaat	been
'veer'	'grasshopper'	tomato	leg
bike /baɪk/	bijl /bɛil/	meisje	wolk
'fiets'	'axe'	girl	cloud
bowl /boʊl/	boom /bom/	auto	snor
'kom'	'tree'	car	moustache
carrot /kærət/	kerk /kɛrk/	fluitje	spiegel
'wortel'	'church'	whistle	mirror
closet /klazɪt/	klomp /klɔmp/	paddestoel	oog
'wandkast'	'wooden shoe'	mushroom	eye
duck /dʌk/	dak /dɑk/	schaar	vliegtuig
'eend'	'roof'	scissors	plane
flashlight /flæʃlaɪt/	fles /flɛs/	klerenhanger	muis
'zaklantaarn'	'bottle'	coat hanger	mouse
knife /naɪf/	nijlpaard /nɛilpart/	schelp	vogelhuis
'mes'	'hippopotamus'	shell	birdhouse
light bulb /laɪt bʌlb/	lijst /lɛist/	strik	vogel
'gloeilamp'	'frame'	bow	bird
money /mʌni/	mand /mɑnt/	den	aardbei
'geld'	'basket'	pine	strawberry
pie /paɪ/	pijl /pɛil/	deur	kous
'taart'	'arrow'	door	stocking
shark /ʃɑrk/	sjaal /ʃal/	föhn	berg
'haai'	'scarf'	hairdryer	mountain
spine /spaɪn/	spijker /spɛikər/	hoed	raam
'ruggegraat'	'nail'	hat	window
stamp /stæmp/	stekker /stɛkər/	ezel	bril
'postzegel'	'plug'	donkey	glasses

Target–competitor pairs with closely approximating vowels are underlined. In Experiment 4, the Dutch competitor words became the target words.

References

Allopenna, P., Magnuson, J., & Tanenhaus, M. (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.

Art explosion library [Computer software] (1995). Calabasas, CA: Nova Development Corporation.

Baayen, H., Piepenbrock, R., & van Rijn, H. (1993). The CELEX lexical database (CD-ROM). Philadelphia: Linguistic Data Consortium, University of Pennsylvania.

Best, C. (1995). A direct realist view of cross-language speech perception: Standing at the crossroads. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 171–204). Baltimore: York Press.

Best, C., McRoberts, G., & Goodell, E. (2001). Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listener's native phonological system. *Journal of the Acoustical Society of America, 109*, 775–794.

Best, C., McRoberts, G., & Sithole, N. (1988). Examination of perceptual reorganization for non-native speech contrasts: Zulu click discrimination by English-speaking adults and

infants. *Journal of Experimental Psychology: Human Perception and Performance, 14*, 345–360.

Best, C., & Strange, W. (1992). Effects of phonological and phonetic factors on cross-language perception of approximants. *Journal of Phonetics, 20*, 305–330.

Bijeljac-Babic, R., Biardeau, A., & Grainger, J. (1997). Masked orthographic priming in bilingual word recognition. *Memory & Cognition, 25*, 447–457.

Broersma, M. (2002). Comprehension of non-native speech: Inaccurate phoneme processing and activation of lexical competitors. In: *Proceedings of the 7th international conference on spoken language processing* (pp. 261–264). Center for Spoken Language Research, University of Colorado, Boulder (CD-ROM).

Cutler, A., Weber, A., Smits, R., & Cooper, N. (forthcoming). Patterns of English phoneme confusions by native and non-native listeners.

Cycowicz, Y., Friedman, D., Rothstein, M., & Snodgrass, J. (1997). Picture naming by young children: Norms for name agreement, familiarity, and visual complexity. *Journal of Experimental Child Psychology, 65*, 171–237.

Dahan, D., Magnuson, J., & Tanenhaus, M. (2001). Time course of frequency effects in spoken-word recognition: Evidence from eye movements. *Cognitive Psychology, 42*, 317–367.

Dahan, D., Swingley, D., Tanenhaus, M., & Magnuson, J. (2000). Linguistic gender and spoken-word recognition in French. *Journal of Memory and Language, 42*, 465–480.

Dijkstra, A., Grainger, J., & van Heuven, W. (1999). Recognition of cognates and interlingual homographs: The neglected role of phonology. *Journal of Memory and Language, 41*, 496–518.

Dijkstra, A., Timmermans, M., & Schriefers, H. (2000). On being blinded by your other language: Effects of task demands and language intermixing. *Journal of Memory and Language, 42*, 445–464.

Dijkstra, A., van Jaarsveld, H., & ten Brinke, S. (1998). Interlingual homograph recognition: Effects of task demands and language intermixing. *Bilingualism: Language & Cognition, 1*, 51–66.

Doctor, E., & Klein, D. (1992). Phonological processing in bilingual word recognition. In R. Harris (Ed.), *Cognitive Processing in Bilinguals* (pp. 237–252). Amsterdam: Elsevier.

Fischer, B. (1992). Saccadic reaction time: Implications for reading, dyslexia and visual cognition. In K. Rayner (Ed.), *Eye Movements and Visual Cognition: Scene Perception and Reading* (pp. 31–45). New York: Springer Verlag.

Flege, J., & Frieda, E. (1997). Amount of native-language (L1) use affects the pronunciation of an L2. *Journal of Phonetics, 25*, 169–186.

Goldinger, S., Luce, P., Pisoni, D., & Marcario, J. (1992). Form-based priming in spoken word recognition: The roles of competition and bias. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*, 1211–1238.

Grosjean, F. (1988). Exploring the recognition of guest words in bilingual speech. *Language and Cognitive Processes, 3*, 233–274.

Guion, S., Flege, J., Akahane-Yamada, R., & Pruitt, J. (2000). An investigation of current models of second language speech perception: The case of Japanese adults' perception of English consonants. *Journal of the Acoustical Society of America, 107*, 2711–2724.

Hillenbrand, J., Getty, L., Clark, M., & Wheeler, K. (1995). Acoustic characteristics of American English vowels. *Journal of the Acoustical Society of America, 97*, 3099–3111.

Ju, M., & Luce, P. (in press). Falling on sensitive ears: Constraints on bilingual lexical activation. *Psychological Science*.

Li, P. (1996). Spoken word recognition of code-switched words by Chinese-English bilinguals. *Journal of Memory and Language, 35*, 757–774.

Longman Dictionary of Contemporary English (2nd ed.). (1987). London: Longman Group UK Limited.

Maddieson, I. (1984). *Patterns of Sounds*. Cambridge: Cambridge University Press.

Marian, V., & Spivey, M. (1999). Activation of Russian and English cohorts during bilingual spoken word recognition. In *Proceedings of the 21st Annual Conference of the Cognitive Science Society* (pp. 349–354). Mahwah, NJ: Erlbaum.

Marian, V., & Spivey, M. (2003). Bilingual and monolingual processing of competing lexical items. *Applied Psycholinguistics, 24*, 173–193.

Marian, V., & Spivey, M., (in press). Competing activation in bilingual language processing: Within- and between-language competition. *Bilingualism: Language and Cognition*.

Marslen-Wilson, W. D., & Welsh, A. (1978). Processing interactions and lexical access during word-recognition in continuous speech. *Cognitive Psychology, 10*, 29–63.

Matin, E., Shao, K., & Boff, K. (1993). Saccadic overhead: Information processing time with and without saccades. *Perception & Psychophysics, 53*, 372–380.

McClelland, J., & Elman, J. (1986). The TRACE model of speech perception. *Cognitive Psychology, 18*, 1–86.

McMurray, B., Tanenhaus, M., Aslin, R., & Spivey, M. (2003). Probabilistic constraint satisfaction at the lexical/phonetic interface: Evidence for gradient effects of within-category VOT on lexical access. *Journal of Psycholinguistic Research, 32*, 77–97.

McQueen, J., Norris, D., & Cutler, A. (1994). Competition in spoken word recognition: Spotting words in other words. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*, 621–638.

Meador, D., Flege, J., & MacKay, I. (2000). Factors affecting the recognition of words in a second language. *Bilingualism, 3*, 55–67.

Nas, G. (1983). Visual word recognition in bilinguals: Evidence for a cooperation between visual and sound based codes during access to a common lexical store. *Journal of Verbal Learning and Verbal Behavior, 22*, 526–534.

Norris, D., McQueen, J., & Cutler, A. (1995). Competition and segmentation in spoken-word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 1209–1228.

Pallier, C., Colomé, A., & Sebastián-Gallés, N. (2001). The influence of native-language phonology on lexical access: Exemplar-based versus abstract lexical entries. *Psychological Science, 12*, 445–449.

Polka, L. (1995). Linguistic influences in adult perception of non-native vowel contrasts. *Journal of the Acoustical Society of America, 97*, 1286–1296.

Salverda, A. P., Dahan, D., & McQueen, J. (in press). The role of prosodic boundaries in the resolution of lexical embedding in speech comprehension. *Cognition*.

Saslow, M. (1967). Latency for saccadic eye movement. *Journal of the Optical Society of America, 57*, 1030–1033.

Schulpen, B., Dijkstra, A., Schriefers, H., & Hasper, M. (in press). Recognition of interlingual homophones in bilingual auditory word recognition. *Journal of Experimental Psychology: Human Perception and Performance*.

Snodgrass, J., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory, 6*, 174–215.

Spivey, M., & Marian, V. (1999). Crosstalk between native and second languages: Partial activation of an irrelevant lexicon. *Psychological Science, 10*, 281–284.

Strange, W. (Ed.). (1995). *Speech perception and linguistic experience: Issues in cross-language research*. Baltimore: York Press.

Tanenhaus, M., Magnuson, J., Dahan, D., & Chambers, C. (2000). Eye movements and lexical access in spoken-language comprehension: Evaluating a linking hypothesis between fixations and linguistic processing. *Journal of Psycholinguistic Research, 29*, 557–580.

Tanenhaus, M., & Spivey-Knowlton, M. (1996). Eye-tracking. *Language and Cognitive Processes, 11*, 583–588.

Tanenhaus, M., Spivey-Knowlton, M., Eberhard, K., & Sedivy, J. (1995). Integration of visual and linguistic information during spoken language comprehension. *Science, 268*, 1632–1634.

van Heuven, W., Dijkstra, A., & Grainger, J. (1997). Shared neighborhood effects in masked orthographic priming. *Psychonomic Bulletin & Review, 8*, 96–101.

Vroomen, J., & de Gelder, B. (1995). Metrical segmentation and lexical inhibition in spoken word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 21*, 98–108.

Zwitserlood, P. (1989). The locus of the effects of sentential-semantic context on spoken-word processing. *Cognition, 32*, 25–64.