

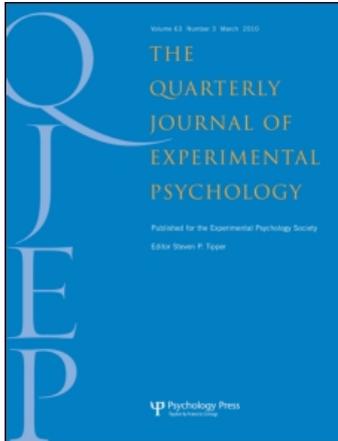
This article was downloaded by: [Max Planck Institute for Psycholinguistic]

On: 6 January 2011

Access details: Access Details: [subscription number 918399536]

Publisher Psychology Press

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



## The Quarterly Journal of Experimental Psychology

Publication details, including instructions for authors and subscription information:

<http://www.informaworld.com/smpp/title~content=t716100704>

### Competition dynamics of second-language listening

Mirjam Broersma<sup>ab</sup>, Anne Cutler<sup>abc</sup>

<sup>a</sup> Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands <sup>b</sup> Radboud University Nijmegen, Donders Institute for Brain, Cognition and Behaviour, Nijmegen, The Netherlands <sup>c</sup> MARCS Auditory Laboratories, University of Western Sydney, Sydney, Australia

First published on: 23 August 2010

**To cite this Article** Broersma, Mirjam and Cutler, Anne(2011) 'Competition dynamics of second-language listening', The Quarterly Journal of Experimental Psychology, 64: 1, 74 – 95, First published on: 23 August 2010 (iFirst)

**To link to this Article:** DOI: 10.1080/17470218.2010.499174

**URL:** <http://dx.doi.org/10.1080/17470218.2010.499174>

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: <http://www.informaworld.com/terms-and-conditions-of-access.pdf>

This article may be used for research, teaching and private study purposes. Any substantial or systematic reproduction, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

# Competition dynamics of second-language listening

Mirjam Broersma<sup>1,2</sup>, and Anne Cutler<sup>1,2,3</sup>

<sup>1</sup>Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands

<sup>2</sup>Radboud University Nijmegen, Donders Institute for Brain, Cognition and Behaviour, Nijmegen, The Netherlands

<sup>3</sup>MARCS Auditory Laboratories, University of Western Sydney, Sydney, Australia

Spoken-word recognition in a nonnative language is particularly difficult where it depends on discrimination between confusable phonemes. Four experiments here examine whether this difficulty is in part due to phantom competition from “near-words” in speech. Dutch listeners confuse English /æ/ and /ɛ/, which could lead to the sequence *daf* being interpreted as *deaf*, or *lemp* being interpreted as *lamp*. In auditory lexical decision, Dutch listeners indeed accepted such near-words as real English words more often than English listeners did. In cross-modal priming, near-words extracted from word or phrase contexts (*daf* from *DAFfodil*, *lemp* from *eviL EMPire*) induced activation of corresponding real words (*deaf*; *lamp*) for Dutch, but again not for English, listeners. Finally, by the end of untruncated carrier words containing embedded words or near-words (*definite*; *daffodil*) no activation of the real embedded forms (*deaf* in *definite*) remained for English or Dutch listeners, but activation of embedded near-words (*deaf* in *daffodil*) did still remain, for Dutch listeners only. Misinterpretation of the initial vowel here favoured the phantom competitor and disfavoured the carrier (lexically represented as containing a different vowel). Thus, near-words compete for recognition and continue competing for longer than actually embedded words; nonnative listening indeed involves phantom competition.

**Keywords:** Word recognition; Competition; Listening; First language; Second language.

The sounds of a second language (L2) can famously cause perceptual difficulty. A substantial body of research findings now documents this issue, across many types of sounds and many pairs of languages (see, e.g., the collected papers in Bohn & Munro, 2007, and Strange, 1995).

However, the goal of speech perception is not the recognition of sounds, but the recognition of words and the meaning they convey. Listeners suffer from the difficulties of sound perception only to the extent that their recognition of words and meaning is affected. How difficulties with

---

Correspondence should be sent to Mirjam Broersma, Max Planck Institute for Psycholinguistics, PO Box 310, 6500 AH Nijmegen, The Netherlands. E-mail: mirjam@mirjambroersma.nl

This research was supported by a VENI grant from the Netherlands Organisation for Scientific Research (NWO) to the first author, and by a SPINOZA award from the NWO to the second author. We would like to thank Antje Meyer of the Behavioral Brain Science Center, University of Birmingham, and Alan Garnham of the Laboratory of Experimental Psychology, University of Sussex, for giving us the opportunity to test participants in their labs. Thanks also to Nicolas Dumay and two anonymous reviewers for useful comments.

the perception of speech sounds play out as problems in L2 listeners' word recognition is the topic of this paper.

Inaccurate sound perception results from mismatches between the phonemic repertoires of the first, or native, language (L1) and a second language acquired post childhood. Such mismatches can occur either with consonants (for example, the well-known difficulty of the English /r/-/l/ contrast for many Asian learners of English; Goto, 1971) or with vowels (for example, the English /æ/-/ɛ/ contrast for listeners with Dutch or German L1; Schouten, 1975). These mismatches can cause lexical confusion in at least three different ways. First, minimal pairs such as *write* and *light* or *flash* and *flesh* can be conflated. Thus Broersma (2002, 2005b) found that even for Dutch listeners with a high level of proficiency in their L2 English, presentation of one word (e.g., *flash*) in a cross-modal priming task led to strong lexical activation of the other word (e.g., *flesh*). Dutch listeners produced similar results in a repetition priming task, as did Japanese listeners, who showed joint activation of pairs like *write*–*light* (Cutler & Otake, 2004), and even highly fluent early Spanish–Catalan bilinguals showed the same for minimal pairs differing in Catalan-only contrasts such as /netə/-/netə/ (Pallier, Colomé, & Sebastián-Gallés, 2001). The inability to tell apart pairs of words with quite different meanings has the potential to lead to a breakdown of communication.

Less obviously, words that are distinct overall, but differ early on in a perceptually difficult sound contrast, can become temporarily indistinguishable. Thus, hearing *panda* temporarily activated *pencil* for Dutch listeners (Weber & Cutler, 2004), and *rocket* activated *locker* for Japanese listeners (Cutler, Weber, & Otake, 2006). Such extended availability of alternative interpretations of speech may not irreversibly disrupt L2 listeners' processing, but it could certainly slow it down.

The third and least obvious problem is that parts of one or more words in the speech signal may resemble a word that is not at all present in the speech. Consider that the word *DAFfodil*

contains the syllable *daf*; for Dutch listeners, this might resemble the word *deaf*. Similarly, the phrase *eviL EMPire* contains *lemp*, which resembles the real word *lamp* in much the same way. Such “near-words” (the upper case portions in the example words) might seem, to L2 listeners, to be viable words and might thus compete for recognition. For L1 listeners they should not play a role in word recognition at all, so that this third possibility could also saddle nonnative listeners with lexical competition over and above that which native listeners have to deal with.

There is experimental evidence that this third type of confusion, too, affects the word recognition process. Fluent Spanish–Catalan bilinguals have been shown not to differentiate well between words and near-words of Catalan such as /finestʁə/-/finestʁə/ (Sebastián-Gallés, Echeverría, & Bosch, 2005). Listeners with Dutch L1, in which consonant voicing is not word-finally distinctive, experience activation of real English words (*groove*, *flight*) when they hear near-words like *groof* and *flide* (Broersma & Cutler, 2008). Note that in the latter case, inability to perceive the speech sound distinction is not at issue, since consonant voicing distinctions do indeed occur in Dutch (e.g., the /t/ and /d/ in *Amsterdam* are distinct sounds); but Dutch does not have them in all the positions that English does. Dutch listeners are actually good at perceiving voicing distinctions in syllable-final position when presented with artificially created nonword continua (Broersma, 2005a, 2008), but because words of their L1 do not distinguish voicing in this position, they overlook the necessity to make such distinctions in real words in L2 speech. These findings suggest that activation of embedded near-words may also be a significant factor in L2 word recognition.

Simultaneous activation of multiple lexical candidates, and subsequent competition between them, is a central assumption of all current models of speech comprehension (for a review see McQueen, 2005). It is held to be, indeed, the basis of the outstanding efficiency of spoken-word recognition. Nevertheless, how much lexical activation and competition occurs is

crucial: The more competitors there are, the slower recognition of an intended word can be (Norris, McQueen, & Cutler, 1995; Vroomen & De Gelder, 1995). Phantom competitors, not actually present in the speech signal but similar to parts of one or more of the words that are present, may therefore provide additional lexical competition over and above the competition that L1 listeners have to deal with. In all the ways described above, word recognition for L2 listeners would then become significantly less efficient.

Moreover, the status of phantom competitors and real competitors may be unequal, and not in the way one might prefer: Phantom competitors may actually compete more strongly than real competitors. Consider that the experimental evidence from L2 listening so far suggests that confusable contrasts such as /r/-/l/ or /æ/-/ɛ/ are initially interpreted as unambiguously one member of the phoneme pair, and words in the lexicon containing that phoneme are preferentially activated (Cutler et al., 2006; Escudero, Hayes-Harb & Mitterer, 2008; Weber & Cutler, 2004). Words containing the other member of the pair are, correspondingly, not activated by the input. In other words, the distinction that is perceptually confused in the speech input is nevertheless correctly represented as distinct in the lexicon. Phantom words may then prove to be more serious competitors than other activated word forms, exactly because their activation is based on an erroneous phonemic interpretation and that erroneous interpretation will not match the correct representation in the lexicon. If a listener misperceives the crucial vowel in, for example, *DAFfodil* as /ɛ/ instead of /æ/, activation of *deaf* may be very persistent, as *deaf* might seem a phonetically more defensible interpretation than *daffodil*.

In this paper, we report a direct investigation of such phantom word activation, using near-words based on the vowel contrast of /æ/ and /ɛ/ (*bat*, *bet*) in British English, presented to Dutch listeners with high proficiency in English. This vowel contrast, as already noted, is difficult for Dutch listeners to distinguish even when their level of English proficiency is high (Broersma, 2005a;

Schouten, 1975). The reason for this is that it is the hardest type of L2 phonetic distinction, one where two categories in the L2 correspond to a single category in the L1 (Best, 1994); in the area of phonetic space in which English has both /æ/ and /ɛ/, Dutch has only one vowel (Booij, 1995). Although this Dutch vowel is represented in the International Phonetic Alphabet (IPA) by /ɛ/, it is not identical to the English vowel represented by the same IPA symbol; rather, its typical acoustic realization falls between the typical realizations of English /æ/ and /ɛ/ (for the detailed acoustic structure of the two vowel systems, based on recent measurements, see Deterding, 1997, for British English, and Adank, van Hout, & Smits, 2004, for Dutch).

We assess the effects of this vowel system mismatch by a series of four experiments. First, a lexical decision study addresses the baseline question of whether Dutch listeners indeed perceive near-words like *lemp* as real words. For this initial experiment, near-words were recorded in isolation. Next, two cross-modal priming studies examine whether there is also phantom word activation when near-words are taken from a context in which they might occur in normal speech. In Experiment 2, near-words are taken from an embedding carrier word, such as *DAFfodil*, with the embedded near-word corresponding to the first syllable of the carrier, while in Experiment 3 they are extracted from a phrase such as *eviL EMPire*, with the near-word occurring in this case across a word boundary. In each case, we check for activation of the real-word counterpart of the near-word (*deaf*, *lamp*, etc.). Finally, the crucial question is addressed, again via cross-modal priming, whether phantom word activation influences word recognition when the stimuli are not isolated from their carriers but encountered as they would be in speech. In Experiment 4, therefore, listeners are presented with complete carriers such as *DAFfodil* while, again, activation of the real word *deaf* is assessed.

In all of the experiments, we compare the Dutch listeners' results to those of native British English listeners presented with the same input. It is obvious—given the fact that the single Dutch

category overlaps the two English categories—that the vowel contrast we use involves two acoustically quite close signals; the contrast might thus also easily be misperceived by L1 listeners (consider, for instance, that Australian listeners performing vowel categorization on American English vowels produced one of their three highest error rates on this distinction; Cutler, Smits, & Cooper, 2005). Further, it is of course the case that nonwords may temporarily activate the nearest existing word for native listeners too (Newman, Sawusch, & Luce, 1997; Taft, 1986). Thus it is important to know whether our nonnative listeners' results are indeed different from the response patterns that native listeners would produce, if we are to use them to motivate conclusions about the nature of nonnative lexical processing.

## EXPERIMENT 1

Experiment 1 assesses whether phantom word activation ensues when near-words are heard. In auditory lexical decision, do Dutch listeners respond “yes” to forms such as *lemp* or *chast* more often than English listeners do?

### Method

#### *Participants*

Participants were 24 native speakers of Dutch and 24 native speakers of British English. The Dutch participants had a high level of proficiency in English as a second language. They had received at least six and on average eight years of English instruction in primary and secondary education and were regularly exposed to English through the media and at the university they attended. The English participants did not know any Dutch. The Dutch participants were recruited from the Max Planck Institute participant pool, the English participants at the University of Birmingham. None reported any hearing loss, visual loss, or reading disability. All were volunteers and received a small fee for participation.

#### *Materials*

As experimental items, 32 monosyllabic English words were selected, 16 containing the vowel /æ/, and 16 containing /ɛ/. For each experimental word, a near-word was formed by replacing /æ/ with /ɛ/ or vice versa (e.g., *lamp* became *lemp*, *chest* became *chast*; see Appendix A). Neither the real words nor the near-words sounded like existing Dutch words. The mean logarithmic lemma frequency per million of the experimental words in the CELEX lexical database of British English (Baayen, Piepenbrock, & Gulikers, 1995) was 2.05.

To conceal the manipulation of the vowels /æ/ and /ɛ/, several types of fillers were included. First, there were another 32 words and 32 near-words formed by, in total, eight different consonant replacements, involving phoneme contrasts that are not linguistically distinctive in word-final position in Dutch (/z/ to /s/ and vice versa, /v/ to /f/ and vice versa, /b/ to /p/ and vice versa, and /d/ to /t/ and vice versa, all in word-final position). Further, there were 68 monosyllabic English filler words and 68 nonwords, formed by replacing either the vowel or the final consonant in a real word with a phoneme that was not expected to be confusable for Dutch listeners, in such a way that the phonotactic constraints of English were not violated.

The materials were recorded by a male native speaker of British English. The speaker read the items one by one, separated by a pause, in a clear citation style. The recording was made in a sound-attenuated booth using a high-quality microphone onto digital audiotape and was downsampled to 16 kHz during transfer to a computer.

#### *Design*

The experimental words were divided into two lists, balanced for frequency, with equal numbers of each vowel in each list. Each participant heard the 16 experimental words from one list in their real-word form and the 16 experimental words from the other list in near-word form. Each participant was also presented with 16 of the words and 16 of the near-words based on the voicing manipulation, and with all of the other 68 filler words and 68

nonwords, so that each heard a total of 100 words and 100 non- and near-words. Items were presented in semirandom order, with the restriction that minimally three other items appeared between two experimental words or two near-words.

### Procedure

Participants were tested one at a time in a quiet room. They received written instructions in their native language, informing them that they were going to hear English words and nonwords. They were asked to press a green response button, labelled “yes”, with their dominant hand if they thought the presented item was an English word, and a red response button, labelled “no”, with their nondominant hand if they thought the item was not an English word. Participants were asked to respond both as fast and as accurately as possible. The experiment started with 10 practice trials and was controlled with NESU (Nijmegen Experiment Set-Up) experimental software. Stimuli were presented binaurally over closed headphones at a comfortable listening level, one at a time. Participants responded by pressing one of the two response buttons on a button box in front of them. No time limit was imposed for responses; presentation of the next item started 800 ms after each button press.

## Results and discussion

In this experiment and all the following, reaction times (RTs) were measured from offset of the auditory stimulus, and outliers were removed from the analyses. Arcsine transformations were applied before analysis of proportions of correct responses or “yes” responses.

Figure 1 displays the percentage of “yes” responses to real words and to near-words by the Dutch and English listeners, respectively. As can be seen, the acoustic closeness between near-words and their real-word counterparts resulted in false-alarm “yes” responses from all participants; however, there was an interaction between native language and condition,  $F_1(1, 46) = 64.34, p < .001$ ;  $F_2(1, 30) = 42.60, p < .001$ , in that significantly more false-alarm responses were made by

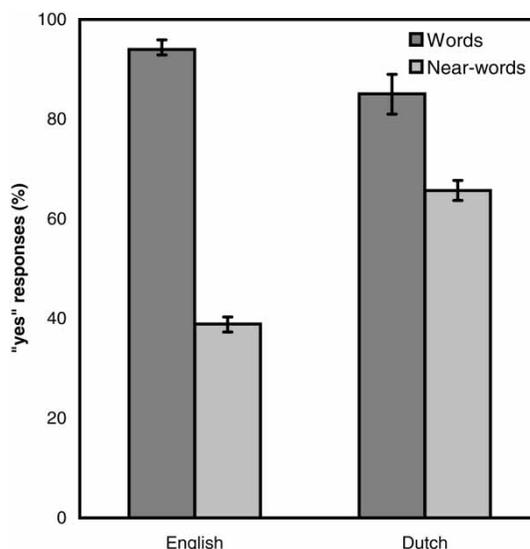


Figure 1. Experiment 1: English and Dutch listeners' percentage of "yes" responses to words and near-words. Error bars indicate 95% confidence intervals.

the Dutch listeners than by the English listeners,  $F_1(1, 46) = 35.09, p < .001$ ;  $F_2(1, 30) = 27.40, p < .001$ . For the real words, on the other hand, the Dutch listeners gave fewer “yes” responses than the English listeners did,  $F_1(1, 46) = 15.12, p < .001$ ;  $F_2(1, 30) = 7.10, p < .05$ . The Dutch listeners did treat real words and near-words differently to some extent, as their “yes” response rate was higher for words than for near-words,  $F_1(1, 23) = 46.46, p < .001$ ;  $F_2(1, 30) = 21.18, p < .001$ , while they did not give more “yes” responses to filler nonwords than the English listeners did (10.1% vs. 11.6%, *ns*).

There was no significant difference in the amount of spurious lexical activation for near-words with an /æ/ and those with an /ɛ/ (for the Dutch listeners, the difference between words and near-words was 15.3% for items like *chest-chast* and 23.5% for items like *lamp-lemp*). Therefore, near-words with an /æ/ and those with an /ɛ/ were collapsed in the above analysis. As the English listeners gave relatively few “yes” responses to near-words, analysis of RTs (see Table 1) is not very informative. Note only that “yes” responses were in general slower to near-words than to words,  $F_1(1, 46) = 20.75, p < .000$ ;  $F_2(1, 56) = 13.88, p < .000$ , and Dutch

**Table 1.** Experiment 1, Lexical decision: English and Dutch listeners' percentage of "yes" responses and reaction times (RTs) of "yes" responses to words and near-words

	Examples	Yes (%)		RT (ms)	
		English	Dutch	English	Dutch
Words	<i>lamp, desk</i>	94.0	85.1	289.3	700.3
Near-words	<i>lemp, dask</i>	38.9	65.7	413.7	781.5

participants responded more slowly than English participants,  $F_1(1, 46) = 79.18, p < .001$ ;  $F_2(1, 56) = 166.96, p < .001$ .

Thus, Experiment 1 has established that near-words are indeed interpreted by Dutch listeners as real English words; this finding then suggests that near-words encountered in natural listening situations could in principle cause substantial phantom word activation for these listeners. In Experiment 2, we further test this hypothesis by assessing whether phantom word activation can be observed when near-words are not recorded in isolation, but are taken from a carrier context in which they might occur in normal speech. It is important to examine embedded forms directly because the wordlikeness of near-words recorded in isolation may not exactly mirror the situation of near-words embedded in longer words; consider that in both English and Dutch, embedded forms such as *dock* in *doctor* and *ham* in *hamster* do not attain the duration expected for the monosyllabic form, and native listeners can use a mismatch with expected duration to rapidly discard the embeddings from the competition process (Davis, Marslen-Wilson, & Gaskell, 2002; Salverda, Dahan, & McQueen, 2003). If nonnative listeners are able to exploit such cross-linguistically consistent durational cues independently of phonemic information, they may be able to reject *deaf* as a word candidate when they hear *daffodil* simply because the first syllable of the longer word is too short to be a good token of *deaf*, even though the perceived vowel quality in that syllable does indeed provide a match to *deaf*.

To measure the activation of *deaf* given *daf* from *daffodil*, we used the paradigm most widely applied in studies of lexical activation: cross-modal priming (Swinney, 1979). This task allows several types of

prime-target relationship; the most suitable version in this case is identity (or form) priming (Zwitserslood, 1996), in which we can directly check for facilitation of responses to a visual word preceded by a spoken form which is an actual utterance of the embedded corresponding near-word (e.g., spoken *daf* preceding visual *deaf*).

## EXPERIMENT 2

Words and near-words were extracted from carrier words in which they occurred as initial embedding. Thus for the word/near-word pair *deaf-daf*, *DEFinite* and *DAFfodil* were recorded, and the monosyllabic word and near-word excised from them (*deaf* from *definite*, *daf* from *daffodil*). These excised forms then served as auditory primes.

If there is activation of embedded forms despite the durational mismatch between monosyllabic words and the same syllables within longer words, then we should expect to observe, for both native and nonnative listeners, facilitated recognition of the target word when it is preceded by an excised form matching to the same word (prime: *def*, target: *deaf*). After a near-word (prime: *daf*, target: *deaf*), less or no facilitation should be observed for the English listeners (cf. Marslen-Wilson, Nix, & Gaskell, 1995), but for the Dutch listeners, phantom word activation, if present, should produce facilitation here as well.

## Method

### Participants

Participants were 36 native speakers of Dutch and 36 native speakers of British English, meeting the description given for Experiment 1. None had

participated in the previous experiment. The Dutch participants were recruited from the Max Planck Institute participant pool and the English participants from the participant pool of the Laboratory of Experimental Psychology of the University of Sussex.

### Materials

As experimental items, 12 monosyllabic English words were selected for visual presentation, six containing /æ/ and six /ɛ/. The mean logarithmic lemma frequency per million of the visual target words in CELEX was 1.44. For each target word, a carrier word was found in which the target word occurred as an initial embedding (e.g., *definite* for *deaf*). For each target word, a mismatch carrier was found, in which the target word almost occurred as an initial embedding, only mismatching in the vowel, such that an /æ/ in the target was an /ɛ/ in the carrier or vice versa (e.g., *daffodil* for *deaf*). Finally, for each target word a phonologically and semantically unrelated word was selected (e.g., *hovercraft* for *deaf*). Half of the carrier, mismatch carrier, and unrelated words were disyllabic, and the other half trisyllabic. The carrier, mismatch carrier, and unrelated words were recorded by the same native speaker of British English and in the same way as described for Experiment 1.

The embedded word was excised from the carrier word to serve as an auditory prime to the visual target in the match condition (e.g., auditory presentation of *def* from *definite*, visual presentation of *deaf*). The embedded near-word was excised from the mismatch carrier word to serve as an auditory prime in the mismatch condition (e.g., auditory presentation of *daf* from *daffodil*, visual presentation of *deaf*). The initial part (the first syllable and sometimes the onset of the second syllable) was excised from the unrelated word to serve as an auditory prime in the control condition (e.g., auditory presentation of *hov* from *hovercraft*, visual presentation of *deaf*). All excisions were made using speech-processing software, with cuts made at a zero crossing in the waveform. All experimental targets, primes and carriers are listed in Appendix B.

Again, to obscure the crucial manipulation of the vowels /æ/ and /ɛ/, different types of fillers were included: 12 visual target words with match, mismatch, and control primes, involving in the mismatch condition the same consonant voicing replacements as those described in Experiment 1, plus three sets of 24 words and 32 nonwords each, for the match, mismatch, and control conditions respectively. The match condition carriers contained the target as an initial embedding (e.g., *FUNdament-fun*; *HIBernate-hibe*). For the mismatch condition, a carrier word was selected that overlapped at onset with the target word except for a vowel (for half of the items) or consonant (for half of the items); in all these cases, the phoneme replacement was expected to be easy to distinguish for the Dutch listeners (e.g., words: *SUPper-sum*; *ITem-oat*; nonwords: *GLAMour-glash*; *DEACon-dake*). The control condition carriers were phonologically and semantically unrelated to the target (e.g., *MERCy-glow*; *INDicate-fub*). Again, all selected words were di- or trisyllabic, and all were recorded and the initial portions excised as for the experimental items. Visually presented items were not spelled like existing Dutch words, and auditorily presented items did not sound like existing Dutch words.

### Design

Each participant saw each experimental visual target only once, with four targets in each of the three conditions: match (preceded by a matching auditory prime), mismatch (preceded by a prime that mismatched the target in the /æ/-/ɛ/ contrast), and control (preceded by a phonologically and semantically unrelated prime). Each participant also received all filler words and nonwords, for a total of 96 words and 96 nonwords, made up of 64 trials in each of the three conditions, and with an equal number of vowel and consonant replacements. Items were presented in semirandom order, such that maximally 5 visually presented words or 5 visually presented nonwords occurred in succession, and any two experimental targets were separated by at least one other item.

### Procedure

The instructions informed participants that they would hear part of an English word, directly after which an English word or nonword would appear on a computer screen. The experiment began with 12 practice trials. Trials consisted of an auditory stimulus with, directly at its offset, presentation of a visual stimulus, which appeared in large font on a computer screen in front of participants. All further aspects of the procedure were as those in Experiment 1.

### Results and discussion

For all experimental items, visual targets were real words, and the correct response was “yes”. In the following analyses of variance (ANOVAs), the dependent variable was the mean RT of correct responses.

The interaction between listener group and condition was significant,  $F_1(2, 140) = 4.74$ ,  $p < .01$ ;  $F_2(2, 22) = 26.27$ ,  $p < .001$ . Figure 2 shows the priming effect (percentage difference from control) for each listener group in the match and mismatch conditions. For the English listeners, it can be seen that there was facilitation in the match condition,  $F_1(1, 35) = 13.58$ ,  $p < .01$ ;  $F_2(1, 11) = 8.02$ ,  $p < .05$ , but not in the mismatch condition, while for the Dutch listeners, there was facilitation in both match and mismatch conditions—match,  $F_1(1, 35) = 11.04$ ,  $p < .01$ ;  $F_2(1, 11) = 8.38$ ,  $p < .05$ ; mismatch,  $F_1(1, 35) = 8.89$ ,  $p < .01$ ;  $F_2(1, 11) = 4.86$ ,  $p < .05$ —and a direct comparison of these two conditions showed that they did not differ (both  $F_1$  and  $F_2 < 1$ ).

Because the accuracy rates in the match and mismatch conditions were lower than those in the control condition, especially for the nonnative listeners, we carried out a further analysis designed to check for a speed–accuracy trade-off.<sup>1</sup> Over a subset of 27 of the 36 Dutch listeners, the mean accuracy in both match and control conditions was 83.0%; for this subset, the mean match RT was 614.8 ms, and the mean control RT

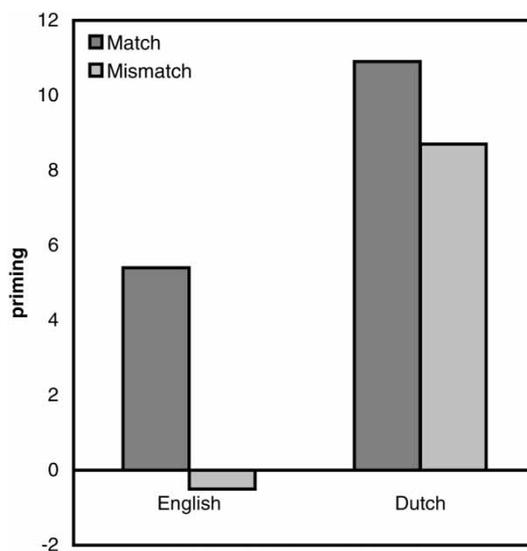


Figure 2. Experiment 2: Priming effects in the match and mismatch conditions, expressed as percentage difference from control (difference between reaction times [RTs] of correct responses in each experimental condition and the control condition, divided by control condition RT,  $\times 100$ ; positive values thus indicate facilitation), separately for English and Dutch listeners.

693.5 ms,  $F_1(1, 26) = 25.3$ ,  $p < .001$ . Over a subset of 29 of these listeners, the mean accuracy in both mismatch and control conditions was 84.2%; for this subset, the mean mismatch RT was 675.1 ms, and the mean control RT 719.7 ms,  $F_1(1, 28) = 5.9$ ,  $p < .05$ . The pattern of RT results cannot therefore be attributed to a speed–accuracy trade-off.

The priming effect from /æ/ to /ɛ/ (*daf-deaf*) was 9.4%, and from /ɛ/ to /æ/ (*ket-cat*) 6.8% for the Dutch listeners, but this was not a significant difference, and near-words with /æ/ and /ɛ/ were collapsed in the above analyses. Not surprisingly, there were again main effects of listener group; the English listeners had shorter correct response RTs than the Dutch listeners,  $F_1(1, 70) = 40.03$ ,  $p < .001$ ;  $F_2(1, 11) = 39.95$ ,  $p < .001$ . [They also had a higher proportion of correct responses:  $F_1(1, 70) = 93.69$ ,  $p < .001$ ;  $F_2(1, 11) = 6.15$ ,

<sup>1</sup> We thank reviewer Nicolas Dumay for this suggestion.

**Table 2.** Experiment 2, Cross-modal priming: English and Dutch listeners' percentage of correct responses and reaction times (RTs) of correct responses for visual target words (e.g., *deaf*) preceded by match, mismatch, and control auditory primes extracted from longer words (e.g., *definite*)

	Examples	Correct (%)		RT (ms)	
		English	Dutch	English	Dutch
Match	<i>cat, deaf</i>	96.4	79.1	534.6	625.8
Mismatch	<i>ket, daf</i>	95.5	79.3	567.8	641.4
Control	<i>pock, hov</i>	97.1	86.3	564.8	702.2

$p < .05$ ; see Table 2. This was the only significant effect in accuracy rate.]

The results of Experiment 2 thus show that near-words excised from naturally spoken carrier words caused phantom word activation for the L2 listeners. For both groups of listeners, recognition of the visual target word was facilitated when it was preceded by auditory presentation of the same word, but, for the Dutch listeners only, recognition of the target word was also facilitated when it was preceded by a near-word. This is particularly interesting in the light of the prior research showing that durational mismatch between monosyllabic words (such as *dock*) and their phonemic matches in longer words (such as *doctor*) reduces activation of such embedded forms (Davis et al., 2002; Salverda et al., 2003). The present findings, also for the native listeners, are compatible with other evidence that shorter words embedded within longer words do indeed become briefly available, though they may rapidly drop out of the set of word candidates due to competition from the longer carrier word (Norris, Cutler, McQueen, & Butterfield, 2006). Certainly it seems clear that the embedded near-words are capable of causing unwanted lexical activation for the nonnative listeners.

Experiment 3 investigated whether phantom word activation is also found for near-words excised from a two-word carrier. Again, it is important to test these excised forms directly, because although cross-word embeddings have been shown to cause lexical activation in native listening (Gow & Gordon, 1995; Tabossi, Burani, & Scott, 1995), there are again durational cues, which on occasion have been shown to reduce activation of

embedded cross-word forms (Shatzman & McQueen, 2006).

### EXPERIMENT 3

The English vocabulary contains a substantial number of almost-embedded words deviating from really embedded words in the /æ/-/ε/ contrast, as Cutler (2005) showed. The number of almost embedded words that might cause phantom word activation is even larger if occurrences across word boundaries are also taken into account. For nonnative listeners, that is, words like *lamp* and *chest* might be activated not only by such real embeddings (e.g., in *eviL AMPlitude* or *eaCH ESTimate*), but also by near embeddings (e.g., in *eviL EMPire* or *eaCH ASTeroid*). To test this, an appropriate two-word carrier context was constructed for each of the stimulus items recorded in isolation in Experiment 1, and for Experiment 3 the same words and near-words were now recorded spanning a word boundary. Thus, the word and near-word pairs *lamp-lemP* and *chest-chast* were presented as forms excised from the carrier contexts *eviL AMPlitude*, *eviL EMPire*, *eaCH ESTimate*, *eaCH ASTeroid*. The task was once again cross-modal identity priming.

### Method

#### Participants

Participants were 36 native speakers of Dutch and 36 native speakers of British English, meeting the

description given in Experiment 2; none had participated in Experiments 1 or 2.

### Materials

The 32 experimental words from Experiment 1 were used as visual targets (see Appendix A). The same real words served as auditory primes for the match condition (e.g., prime: *lamp*; target: *lamp*), and the 32 corresponding near-words as auditory primes for the mismatch condition (e.g., prime: *lemp*; target: *lamp*). Additionally, 32 monosyllabic items, half of them words and half nonwords, all phonologically and semantically unrelated to the target, were selected as control auditory primes (e.g., prime: *bike*; target: *lamp*).

For all match, mismatch, and control primes, a two-word phrase was found in which the desired prime occurred as an embedding across the word boundary, with the first one or two phonemes of the auditory prime in the first word and the rest in the second (e.g., *lamp* from *evil amplitude*, *lemp* from *evil empire*, *bike* from *prefab icon*). In about one third of the cases, auditory prime offset coincided with the offset of the second carrier word (e.g., *span* in *glass pan*; see Appendix A). The phoneme environments of the match and the mismatch primes were kept as similar as possible. For three experimental pairs, no suitable carrier fragments could be found, so fictitious geographic names were created for both match and mismatch primes (these items are marked in Appendix A).

Several types of fillers again concealed the experimental manipulation. The 32 words and 32 near-words of Experiment 1 with consonant voicing changes were augmented by 32 new unrelated words to form a set of 32 filler targets with three types of prime, and 65 further monosyllabic words were selected as filler visual targets, again to be presented with three types of prime. As the match primes here were always words and the mismatch primes nonwords, half of the control primes were words, the other half nonwords. Finally, 129 monosyllabic filler nonwords were chosen as visual targets, 43 with each type of prime. For the filler nonwords, the match primes were always nonwords, the mismatch primes were words, and half of the control primes were words, the other half

nonwords. Half of the filler mismatch primes differed from the corresponding targets in the vowel, and the other half in the final consonant, in all cases involving contrasts that were expected to be easy to distinguish for Dutch listeners. For all primes for the filler word and nonword targets, a two-word fragment was selected in which the carrier occurred as an embedding, similar to the carrier fragments for the experimental items.

The carrier fragments were recorded by the same native speaker of British English and in the same way as described for Experiment 1. Recordings were stored directly to disc at a sample rate of 16 kHz. Auditory primes were again extracted from the carrier fragments.

### Design and procedure

Each participant saw each of the 32 experimental targets only once. The experimental targets were assigned to the match, mismatch, and control conditions such that two conditions contained 11 experimental targets and the third 10, counterbalancing over participants. Each participant was also presented with all filler words and nonwords, distributed over the conditions such that in total each participant received 43 (experimental and filler) words and 43 nonwords in each condition. Items were presented in a semirandom order, as described for Experiment 2. The procedure was the same as that for Experiment 2.

### Results and discussion

Figure 3 again shows the priming effect per condition. Again, the interaction between native language and condition was significant,  $F_1(2, 138) = 4.42, p < .05$ ;  $F_2(2, 62) = 3.18, p < .05$ . For English listeners, there was once more significant facilitation for match trials,  $F_1(1, 35) = 7.80, p < .01$ ;  $F_2(1, 31) = 10.35, p < .01$ , but no significant effect for mismatch trials, while for Dutch listeners, facilitation again appeared in both conditions—match:  $F_1(1, 34) = 11.60, p < .01$ ;  $F_2(1, 31) = 7.93, p < .01$ ; mismatch:  $F_1(1, 34) = 17.91, p < .001$ ;  $F_2(1, 31) = 12.35, p < .001$ —and these two conditions again did not differ (both  $F_1$  and  $F_2 < 1$ ).

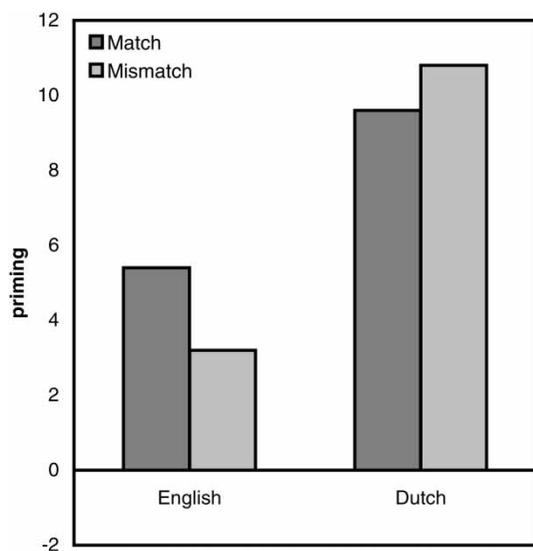


Figure 3. Experiment 3: English and Dutch listeners' priming effects in the match and mismatch conditions, expressed as percentage difference from control.

The priming effect from /æ/ to /ɛ/ was 7.8% and from /ɛ/ to /æ/ 14.1% for the Dutch listeners; this was not significantly different (and in the opposite direction from the difference with the same materials in Experiment 1). As in Experiment 2, the English listeners' correct response RTs were shorter,  $F_1(1, 69)=10.70$ ,  $p < .01$ ;  $F_2(1, 30)=54.11$ ,  $p < .001$ , and their correct response proportion higher,  $F_1(1, 70)=62.50$ ,  $p < .001$ ;  $F_2(1, 30)=31.76$ ,  $p < .001$ , than those of the Dutch listeners (see Table 3; again, the latter effect was the only one in accuracy rate).

Near-words as produced in a two-word context thus induce phantom lexical activation for L2

listeners. Experiment 4 addresses the crucial question of whether such activation also ensues when near-words are encountered within carriers, as they would be in natural speech.

## EXPERIMENT 4

In normal speech, listeners rarely encounter truncated words. Although embedded words are plentiful, ambiguity due to their activation is only temporary, since it will normally be resolved by the disambiguating continuation of the word (Zwitserslood, 1989). Thus, for L1 listeners, who efficiently deactivate competitors after disambiguation, hearing the second and third syllables of *definite* make it clear that the intended word is not *deaf*, and this will lead to deactivation of *deaf*. Activation of the carrier *definite* will be strong and will efficiently inhibit activation of the embedded form.

Whether this situation also obtains for listeners hearing their L2 is now addressed in Experiment 4. Recall that Weber and Cutler (2004) demonstrated that unnecessary multiple activation occurred when Dutch listeners heard, for example, *panda* and initially also activated *pencil*. But Weber and Cutler did not examine whether the embedded words *pan* and *pen* also stayed active for longer than they would have for L1 listeners. That is the case we address here. In principle, the L2 listeners could use the second and third syllables of the word *daffodil* to deactivate the phantom competitor *deaf*, in exactly the way they would with embedded real competitors in their L1. In that case, the presence of embedded

Table 3. Experiment 3, Cross-modal priming: English and Dutch listeners' percentage of correct responses and reaction times (RTs) of correct responses for visual target words (e.g., *lamp*) preceded by match, mismatch, and control auditory primes extracted from cross-word contexts (e.g., *evil amplitude*)

	Examples	Correct (%)		RT (ms)	
		English	Dutch	English	Dutch
Match	<i>lamp, desk</i>	97.7	87.0	557.2	633.7
Mismatch	<i>lemp, dask</i>	96.6	83.9	570.2	624.8
Control	<i>bike, sun</i>	94.3	84.3	588.9	700.6

near-words would only induce a temporary addition of activation, augmenting the temporary multiple activation due to real embeddings, but perhaps not significantly impinging upon L2 listening efficiency. However, as argued in the introduction, phantom competitors may actually compete more strongly than real competitors if the actually uttered vowel is interpreted as the one contained in the phantom competitor rather than the one in the carrier. In that case, phantom competition could seriously disrupt L2 listening.

In the previous experiments, in which near-word stimuli were presented in isolation, there were no strong lexical competitors beyond the paired real words (e.g., *deaf*). Thus, if the word were indeed activated by the input, it would not incur serious competition. In Experiment 4, however, we present the full carrier word (e.g., *daffodil*), which is a better match with the total input than the almost-embedded word *deaf* is. This allows us to examine how almost-embedded words participate in competition with their carriers.

## Method

Experiment 4 was identical to Experiment 2, except that the auditory primes were not the excised portions, but the full carrier words (see Appendix B). Thus, for example, the visual target *deaf* was preceded by *definite* in the match condition, by *daffodil* in the mismatch condition, and by *hovercraft* in the control condition, each word untruncated. Since the prime-target relationship is here between full words, we report the frequencies: Mean log frequency for match primes was 1.1, for mismatch primes 1.31, and for control primes 1.15; neither these frequencies nor the frequency difference between primes and visual targets across prime conditions differed significantly (all  $F_s < 1$ ).

Again, 36 native speakers of Dutch and 36 native speakers of British English took part, all meeting the description given in Experiment 1, and none having participated in Experiments 1 to 3. Materials, design, and procedure were the same as those in Experiment 2.

## Results and discussion

Priming effects per condition are displayed in Figure 4. It can be seen that there are clear differences between the Dutch and English listeners' response patterns. Whereas the RTs in all three conditions were virtually identical for the English listeners (see also Table 4), this was not the case for the Dutch listeners. The Dutch group showed a significant main effect of condition,  $F_1(2, 66) = 5.41, p < .01$ ;  $F_2(2, 22) = 3.65, p < .05$ , while no such effect appeared for the English group. The effect of condition for the Dutch listeners was, however, not symmetrical; facilitation in the match condition in fact failed to reach significance, while facilitation in the mismatch condition was significant across both subjects and items,  $F_1(1, 34) = 8.56, p < .01$ ;  $F_2(1, 11) = 7.08, p < .05$ , and the intercondition comparison was here significant,  $F_1(1, 33) = 6.7, p < .05$ ;  $F_2(1, 11) = 4.89, p < .05$ .

The priming effect for the Dutch listeners was 15.3% from /æ/ to /ɛ/ (*daffodil-deaf*) and 0.9% from /ɛ/ to /æ/ (*kettle-cat*), but this was not significantly different. English listeners had both shorter correct response RTs,  $F_1(1, 68) = 14.67$ ,

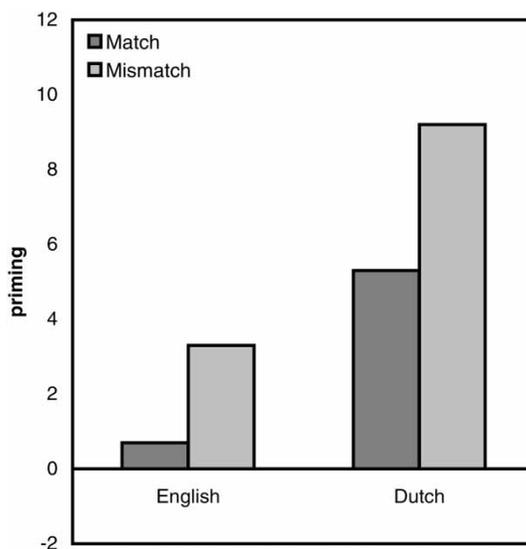


Figure 4. Experiment 4: English and Dutch listeners' priming effects in the match and mismatch conditions, expressed as percentage difference from control.

**Table 4.** Experiment 4, Cross-modal priming: English and Dutch listeners' percentage of correct responses and reaction times (RTs) of correct responses for visual target words (e.g., *deaf*) preceded by full-word match (*definite*), mismatch (*daffodil*), and control (*bovercraft*) auditory primes

	Examples	Correct (%)		RT (ms)	
		English	Dutch	English	Dutch
Match	<i>cat, deaf</i>	97.8	72.9	616.3	703.3
Mismatch	<i>ket, daf</i>	96.4	81.6	600.2	674.4
Control	<i>pock, hov</i>	99.3	78.5	620.5	742.5

$p < .001$ ;  $F_2(1, 11) = 45.78$ ,  $p < .001$ , and more correct responses,  $F_1(1, 70) = 79.20$ ,  $p < .001$ ;  $F_2(1, 11) = 11.58$ ,  $p < .01$ , than the Dutch listeners (see Table 4; again, this latter effect was the only one in accuracy rate).

The results thus show that the L1 listeners here experienced no remaining activation of an embedded word after presentation of the entire word; and of course they experienced little problem at all from near-words because they can accurately perceive the phonemes of their language. This pattern is as expected from previous results. The situation with L2 listeners proved different, but only in a single respect. Their results with really embedded words mimicked those of the L1 listeners, in that little activation remained of the embedded word after presentation of the entire prime—so, after *definite* had been fully heard, *deaf* was no longer effectively competing with it. But their results with spuriously embedded near-words were quite different: In this case, activation of the embedded form did remain—after *daffodil* had been fully heard, *deaf* was still measurably activated. Misinterpretation of the input (in this case, a single vowel) thus has far-reaching consequences: Words matching the erroneous interpretation are activated, and because they are not really embedded in the carrier at all, they cannot be dispatched by competition in the way that real embeddings are so efficiently dealt with.

## GENERAL DISCUSSION

Misperception leads to phantom word activation in nonnative listening: Near-words like *lemp* and

*daf* caused more lexical activation for L2 listeners to English than for native listeners. In Experiment 1, Dutch listeners accepted such near-words as real words more often than English listeners did. In Experiments 2 and 3, near-words extracted from a single word or a two-word carrier facilitated recognition of the words they resembled, but again only for Dutch listeners, not for L1 listeners. In Experiment 4, finally, L2 listeners experienced phantom competition given carriers with near-word embeddings; L1 listeners again did not.

In Experiments 1 to 3, in the absence of a lexical competitor, the paired word was presumably the best match for the near-word input, regardless of the exact interpretation of the vowel. Nonetheless, this situation did not lead L1 listeners to show evidence of significant activation in the priming experiments, and their false-alarm rate in the lexical decision experiment was also significantly lower than that of the L2 listeners. The structure of vocabularies—tens or hundreds of thousands of words, constructed from just a few dozen phonemes—makes embedding inevitable, and listeners are well attuned to the phoneme contrasts of their native language and, in a native listening situation, are accustomed to rejecting word candidates that are not really an accurate match. L2 listeners, no matter how efficient their listening in their own language, cannot achieve the same level of efficiency if the phonemes of the L2 remain confusable for them. In Experiments 1 to 3, not only did the L2 listeners often misinterpret the vowel, this misinterpretation then led them to accept the near-word as a real word in lexical decision and to activate that word in the priming studies.

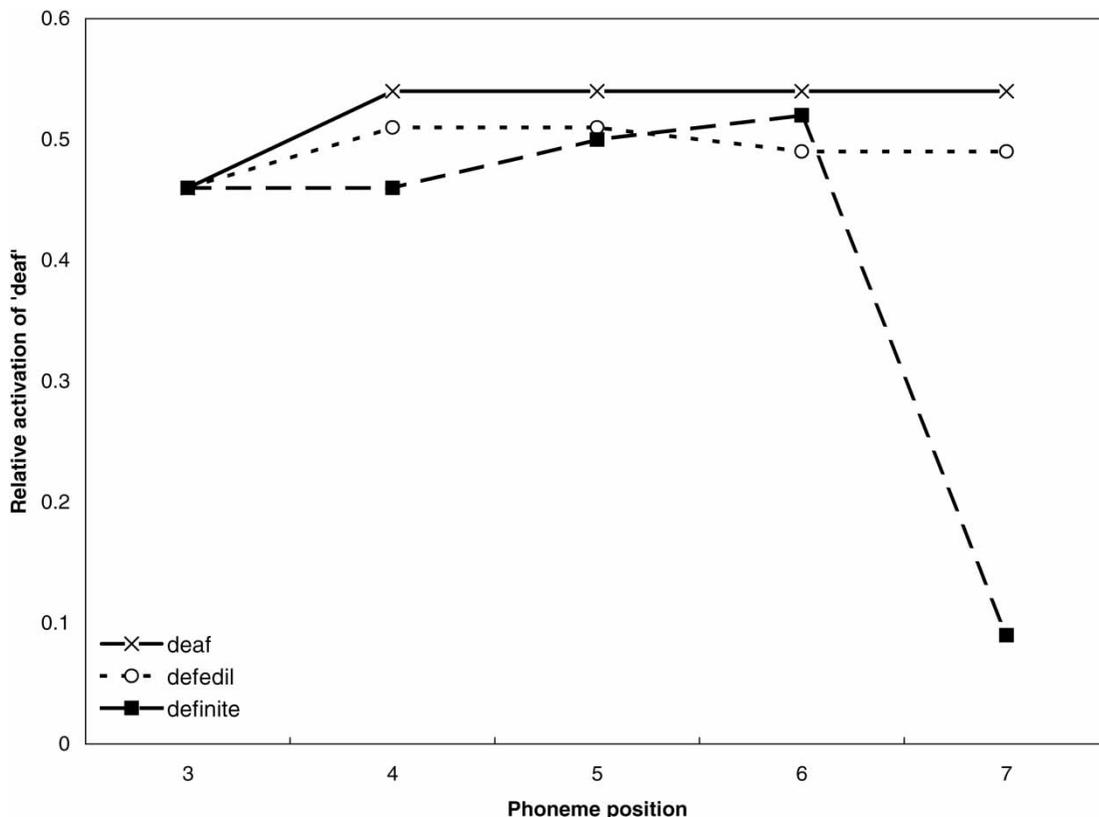
The results of Experiments 1 to 3, though they are based on stimuli presented out of context, provide the framework within which we can interpret L2 listeners' perception of longer stretches of speech as well. Experiment 4 addressed this perceptual case. In this experiment, competition played a role: The embedded near-word, if activated, would have to compete with the carrier word. In this case, a correct interpretation of the vowel would favour the carrier, while an incorrect interpretation would favour the embedded near-word. The fact that the carrier did not always deactivate the nearly embedded word shows that the Dutch listeners misinterpreted the vowel at least some of the time. It also shows that even the presence of a competitor that matches a larger part of the input does not solve the problem of phantom word activation. Misperception of the vowel apparently made the phantom word a strong lexical candidate.

To further explore the competition among lexical candidates, and in particular the relative availability of words like the targets of Experiment 4, we conducted simulations with a computational spoken-word recognition model, Shortlist (Norris, 1994; Norris, McQueen, Cutler, & Butterfield, 1997). Shortlist is a connectionist model based on simultaneous activation and interword competition, and it runs on a realistically sized English dictionary (26,000 words). In the model, matching input increases activation of any word that it matches, mismatching input decreases activation of any existing candidate that it mismatches, and words not only accrue activation but, in proportion to their level of activation, exercise inhibition on other words that are competing for any part of their constituent input.

In four separate simulations, we observed the set of activated candidates when the model was presented with particular inputs. In one simulation, we presented just the three-phoneme word [dɛf] *deaf*, followed by four segments of silence. In the second simulation, we presented the seven-phoneme word [dɛfɪnɪt] *definite*, in the third, the seven-phoneme word [dɛfədɪl] *daffodil*, and in the final simulation [dɛfədɪl] *defedil*—that is, *daffodil* with the vowel in the first syllable replaced by /ɛ/. We cannot directly simulate

differences in listeners' perception, but this set of simulations allows us to compare the crucial cases in our studies: real words versus near-words, and native listeners versus L2 listeners. Because the evidence (Cutler et al., 2006; Weber & Cutler, 2004) indicates that the lexical representations of L2 listeners incorporate veridical contrasts even where their speech perception fails to discriminate the same contrasts, we keep the lexical representations in the model constant; changes in the input then suffice to simulate differences in listeners' perception. Native listeners identify English phonemes correctly; therefore the simulations with *deaf*, *definite*, and *daffodil* all represent native performance. The Dutch listeners identify all these words as having the same vowel in the first syllable; therefore the simulations with *deaf*, *definite*, and *defedil* represent the Dutch performance.

In the case of *daffodil*, the activated set of lexical candidates did not include *deaf* at all. This indeed corresponds to the native listeners' results in Experiment 4; hearing *daffodil* did not lead to facilitation of *deaf*. In all three of the other candidate sets, however, *deaf* was present. This word's relative activation (in the model's activation units, which have no external reference but allow comparison across simulations) is plotted in Figure 5 for each of the three sets. First, the input *deaf* itself activates *deaf*, in accord with the results for both native and L2 listeners hearing *deaf* or *def*- in Experiments 1 to 3. Second, the input *definite* activates *deaf* at first, but then knocks it out of the competition once the seventh phoneme of the carrier word provides overwhelmingly strong evidence that what is being heard is the carrier and not the embedded form. In this latter simulation, the word *definite* itself enters the competitor set at the fifth phoneme, *defin-*, but *in* is also in the set, so that *deaf* and *in*, in combination, for the time being offer enough competition to keep *definite* from triumphing (consider that the utterance might have been *I'd go deaf in a disco*). Only when *definite* is complete to its final phoneme does it amass enough activation to exert inhibition on all competing forms. This is in accord with the results for both native and L2 listeners hearing *definite* in Experiment 4.



**Figure 5.** *The relative activation of the lexical candidate deaf given different input in three separate simulations with Shortlist (Norris et al., 1997). Each input began [def], and each line shows activation from the third input segment, [f], to the seventh. The solid line represents the input deaf followed by silence. It can be seen that deaf is highly activated, rises a little with the first segment of silence (because all competitors, including the closest competitor deft, are mismatched), and stays high. The dashed line shows activation given the input definite. With this input deaf rises less rapidly at the fourth phoneme, because the total competitor set is larger (e.g., deficit, deference, etc.). However, deaf remains available until, at the seventh phoneme, [t], the evidence for definite becomes definite; competition from definite then removes deaf from the candidate set. Either of these two simulations could represent processing by native or by nonnative listeners. The dotted line represents the input defeditil—that is, the way daffodil would be processed if the first vowel were miscategorized. This represents Dutch listeners' processing of daffodil. Crucially, daffodil itself offers no effective competition, because its lexical representation does not match the vowel perceived in the input; thus deaf remains active through the seventh and final phoneme. (N.B., native listeners' processing of daffodil produces no activation of deaf at any point.)*

Third, the particular case of the L2 listeners' misinterpretation of /æ/ as /ɛ/ is addressed in the simulation with the input *defeditil*, and in this case, exactly the same happened as in Experiment 4: The target word *deaf* was activated and remained so, because nothing else was sufficiently activated to succeed in knocking it out of the set. The word *daffodil* was a member of the set by the end of the simulation, but because it was partly mismatched by this deformed input, it

was not strongly activated; indeed, most of the candidate set consisted of such weakly activated, partially mismatched candidates (e.g., *deft*, *ado*, *dildo*), leaving *deaf* to maintain its activation because (a) it was fully matched by the input, and (b) no other fully matched word competed for any part of it.

Computational models such as Shortlist provide an instantiation of the now extensive empirical evidence for simultaneous activation of

word candidates and active competition between them (see McQueen, 2005, for a review). These simulations clearly show the role of competition in our results. When phonemic perception leads to a correct interpretation of all sounds, really embedded forms such as *deaf* in *definite* are efficiently dispatched as a result of the greater competitive contribution of the carrier word, while acoustically similar nearly-embedded forms such as *deaf* given (correctly perceived) *daffodil* do not delay the word recognition process at all. When the interpretation of the initial vowel is incorrect for the word, however (e.g., *daffodil* is interpreted as *defedit*), the nearly embedded word (*deaf*) is bereft of the competitors that should properly have knocked it out. The entire string does not match the stored representation of *daffodil*, because that lexical representation contains a vowel other than the one that was derived from the input. The misinterpretation of the initial vowel thus favours the phantom competitor and disfavours the intended carrier, leading to the result observed both in the simulation and in Experiment 4: Even after all of *daffodil* had been heard, *deaf* remained activated.

Misinterpretation of confusable phonemes does indeed activate word candidates, and these candidates participate in the lexical competition process. Comparisons of the different forms of cross-modal priming have shown that lexical competition involves phonological word forms only (Norris et al., 2006); the degree to which an activated word can hold its own in competition is determined solely by the extent to which it matches or mismatches the phonological interpretation of the speech input. As Experiment 4 showed, embedded near-words hold their own disturbingly well.

Listening to a second language is often experienced as a particularly fragile component of language competence. Highly proficient language users who experience little difficulty with reading, writing, or speaking in their L2 can nevertheless complain that listening rapidly becomes difficult in a noisy environment and that adaptation to unfamiliar talkers and accents falls far short of the automaticity that marks such

adjustment in the L1. Investigations of the effects of noise on listening have revealed that identification of English phonemes is disrupted to an equivalent extent for English L1 and Dutch L2 listeners (Cutler, Weber, Smits, & Cooper, 2004), suggesting that the major difference between L1 and L2 listening is not located at the lowest level of speech perception, but at the higher levels at which compensation for low-level disruption must take place. L1 listeners, in short, recover from disruption far better than L2 listeners. Other findings support this conclusion, by showing better exploitation of sentence predictability by L1 than L2 listeners (Mayo, Florentine, & Buus, 1997) and a correlation between L2 listeners' ability to exploit predictability and their ability to perceive speech in noise (Van Wijngaarden, Steeneken, & Houtgast, 2002). At many levels of speech recognition, L2 listeners' processing falls behind that of L1 listeners, and this makes their listening fragile when any disruption occurs, and recovery is required.

Our study reveals one such clear disadvantage in L2 listening, located at the level of lexical competition. Competition is in essence a highly rapid and efficient process. For L1 listeners, a mismatch between speech signal and phonological representation in the lexicon leads to almost immediate deactivation of the lexical candidate (Soto-Faraco, Sebastián-Gallés & Cutler, 2001; Zwitserlood, 1989), and any traces of inhibition also disappear quickly (McQueen, Norris, & Cutler, 1994). For L2 listeners, there is no reason to expect that the competition process itself will work any differently, and in particular any less efficiently, in a second language from how it does in their first language. But the lexical candidates participating in the competition will not be the same set activated for L1 listeners presented with the same input. If there are more competitors than L1 listeners have to deal with, then L2 listeners will be disadvantaged, as an increase in the competitor population is known to slow spoken-word recognition (Norris et al., 1995; Vroomen & De Gelder, 1995).

An increase in the competitor set for L2 listeners has been reported in other studies; for instance,

Rüschmeyer, Nojack, and Limbach (2008) reported conceptual activation of rhyme competitors for L2 but not for L1 listeners, and evidence that L2 listeners are more sensitive than L1 listeners to the size of a competitor population has come from studies by Bradlow and Pisoni (1999) and by Marian, Blumenfeld, and Boukrina (2008). But our results from the near-embedding case suggest that the situation is actually more complex than a simple increase in competitor set size.

First, our findings confirm that the L2 vocabulary contains distinctions that the L2 perceptual processing cannot take advantage of. As described earlier, Weber and Cutler (2004) and Cutler et al. (2006) showed that confusable sounds such as /t/-/l/ or /æ/-/ɛ/ are often conflated to a single phonemic candidate in perception, but are distinguished in lexical representations (the putative reason for this is that lexical representations learned by L2 users are moulded not only by speech evidence but also by evidence from written text, and by explicit instruction from teachers, textbooks, and native-speaker friends and colleagues; Escudero et al., 2008, provided confirming evidence for the role of such information). In the present study, too, evidence for this appeared; *daffodil* and *definite* did not affect recognition of *deaf* in the same manner, as would have been the case if the representations with which they are stored in the mental lexicon had contained an identical initial vowel. After *def-* had been heard (Experiment 2), *deaf* was activated; these vowels were generally treated as the same. When *definite* was heard (Experiment 4), significant activation of *deaf* had disappeared by the end of the carrier word, just as had been observed with native speakers; again, this suggests that the vowels were treated as the same. When *daffodil* had been heard (Experiment 4), *deaf* remained activated after the end of the carrier, suggesting that the vowels were treated as different—specifically, the vowel in the spoken word *daffodil* is perceived as /ɛ/, leading to activation of *deaf*, but the lexical representation of *daffodil* is stored with another vowel, and hence in competition it is mismatched by the percept.

Second, the crucial role of such misperception is underlined by comparison of the stimuli used here with other types of near-words. When phonemes are not difficult to distinguish, activation of phantom competitors occurs, but those phantom words can be efficiently deactivated in competition. Obstruent voicing contrasts similar to the English contrasts occur in Dutch, but not in word-final position, where only voiceless obstruents can occur. Dutch listeners weight perceptual cues to final voicing (in particular the duration of the preceding vowel) in a different way from how English listeners do, but nevertheless recognize the final English consonants accurately (Broersma, 2005a, 2008). Thus, Dutch listeners should be able to distinguish *groof* from *groove*, and *prez* from *press*, even though the contrast is never relevant in their native language in this position. Broersma and Cutler (2008) tested Dutch and English listeners' perception of near-words such as *groof*, extracted from contexts such as *biG ROOF*. In the absence of lexical competitors, Dutch listeners accepted the near-word as a word more often than English listeners did in lexical decision, and in cross-modal priming, the near-words facilitated recognition of the real words for Dutch listeners only. Thus, the results when lexical competitors played no role were similar to those found in the present study. However, adding lexical competition affected the results differently from the case of the perceptually confusable /æ/-/ɛ/ pairs reported here. Broersma (2007) presented Dutch listeners with full carriers (e.g., prime: *president*, target: *press*) and found no activation of the near-words; the competitor *president* here triumphed over the near-word *press*. Thus, near-words that are perceptually easily distinguishable from real words (Dutch listeners can distinguish voicing contrasts) caused lexical activation when there was no other lexical candidate available (the final voicing of *prez* is not weighted highly enough to mismatch *press*), but not when there was such an alternative (*president* is well supported by the input). Those results contrast with the current findings, where phonetic misinterpretation favoured the nearly embedded word over the actually presented carrier.

It is, therefore, crucially misperception, supported by the lexical accuracy, that induces the phantom competition. If the lexical representation were as inaccurate as the perception, there would be less problem for listeners, because the vowel in the input and the vowel in the lexical representation would be the same, and the competition process would thus be able to dispatch the embedded competitor. Our results show, however, that wherever confusable phonemes occur in the speech input and match a nearly embedded word, phantom competitors will cause problems for the L2 listener. These phantom competitors can be words nearly embedded in carriers (such as *deaf* in *daffodil*), but they could also, we suggest, be words nearly embedded across word boundaries (such as *lamp* in *evil empire*). In the present study, we did not test the latter case in a competition situation, since measurable activation is unlikely to be observed given the availability of an unchallenged candidate for the beginning of the spurious embedding; *evil* will have taken the /l/ that the later occurring candidate *lamp* would need to survive the competition process. In running speech, however, listeners cannot know in advance where words begin; there, embedded words are more likely to be activated (Tabossi et al., 1995), and the same will be true for nearly embedded words too. In normal speech *evil empire* may thus activate *lamp* as a possible competitor for Dutch listeners.

Our study has revealed that phantom word activation arises due to misperception, even when lexical representations are accurate. The misperception can lead to the activation of lexical competitors that are not active for native listeners. Lexical representations that are activated due to such misperceptions are particularly resistant to deactivation, as there is no other lexical candidate in the speech signal that matches the misperception. While the activation of really embedded words does not pose a long-lived problem, for L2 any more than for L1 listeners, the activation of embedded near-words is indeed a potential problem. Phantom competition may seriously hinder nonnative word recognition.

## REFERENCES

- Adank, P., van Hout, R., & Smits, R. (2004). An acoustic description of the vowels of Northern and Southern Standard Dutch. *Journal of the Acoustical Society of America*, *116*, 1729–1738.
- Baayen, H., Piepenbrock, R., & Gulikers, L. (1995). The CELEX lexical database [CD-ROM]. Philadelphia: University of Pennsylvania, Linguistic Data Consortium.
- Best, C. T. (1994). The emergence of native-language phonological influences in infants: A perceptual assimilation model. In J. C. Goodman & H. C. Nusbaum (Eds.), *The development of speech perception: The transition from speech sounds to spoken words* (pp. 167–224). Cambridge, MA: MIT.
- Bohn, O.-S., & Munro, M.J. (Eds.). (2007). *Language experience in second language speech learning: In honor of James Emil Flege*. Amsterdam: John Benjamins.
- Booij, G. (1995). *The phonology of Dutch*. Oxford, UK: Oxford University Press.
- Bradlow, A. R., & Pisoni, D. B. (1999). Recognition of spoken words by native and non-native listeners: Talker-, listener-, and item-related factors. *Journal of the Acoustical Society of America*, *106*, 2074–2085.
- Broersma, M. (2002). Comprehension of non-native speech: Inaccurate phoneme processing and activation of lexical competitors. In J. H. L. Hansen & B. Pellom (Eds.), *Proceedings of the 7th International Conference on Spoken Language Processing* (pp. 261–264). [CD-ROM]. Boulder, CO: University of Colorado, Center for Spoken Language Research.
- Broersma, M. (2005a). Perception of familiar contrasts in unfamiliar positions. *Journal of the Acoustical Society of America*, *117*, 3890–3901.
- Broersma, M. (2005b). *Phonetic and lexical processing in a second language* (MPI Series in Psycholinguistics No. 34). Doctoral dissertation, Radboud University, Nijmegen, The Netherlands.
- Broersma, M. (2007). Why the ‘president’ does not excite the ‘press’: The limits of spurious lexical activation in L2 listening. In J. Trouvain & W. J. Barry (Eds.), *Proceedings of the 16th International Congress of Phonetic Sciences (ICPhS 2007)*, Saarbrücken, Germany [CD-ROM], pp. 1909–1912.

- Broersma, M. (2008). Flexible cue use in nonnative phonetic categorization. *Journal of the Acoustical Society of America*, *124*, 712–715.
- Broersma, M., & Cutler, A. (2008). Phantom word activation in L2. *System: An International Journal of Educational Technology and Applied Linguistics*, *36*, 22–34.
- Cutler, A. (2005). The lexical statistics of word recognition problems caused by L2 phonetic confusion. In *Proceedings of the 9th European Conference on Speech Communication and Technology*. Bonn, Germany: International Speech Communication Association.
- Cutler, A., & Otake, T. (2004, May). *Pseudo-homophony in non-native listening*. Paper presented to the 174th meeting of the Acoustical Society of America, New York.
- Cutler, A., Smits, R., & Cooper, N. (2005). Vowel perception: Effects of non-native language vs. non-native dialect. *Speech Communication*, *47*, 32–42.
- Cutler, A., Weber, A., & Otake, T. (2006). Asymmetric mapping from phonetic to lexical representations in second-language listening. *Journal of Phonetics*, *34*, 269–284.
- Cutler, A., Weber, A., Smits, R., & Cooper, N. (2004). Patterns of English phoneme confusion by native and non-native listeners. *Journal of the Acoustical Society of America*, *116*, 3668–3678.
- Davis, M. H., Marslen-Wilson, W. D., & Gaskell, M. G. (2002). Leading up the lexical garden path: Segmentation and ambiguity in spoken word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, *28*, 218–244.
- Deterding, D. (1997). The formants of monophthong vowels in Standard Southern British English pronunciation. *Journal of the International Phonetic Association*, *27*, 47–55.
- Escudero, P., Hayes-Harb, R., & Mitterer, H. (2008). Novel second-language words and asymmetric lexical access. *Journal of Phonetics*, *36*, 345–360.
- Goto, H. (1971). Auditory perception by normal Japanese adults of the sounds “l” and “r”. *Neuropsychologia*, *9*, 317–323.
- Gow, D. W., & Gordon, P. C. (1995). Lexical and prelexical influences on word segmentation: Evidence from priming. *Journal of Experimental Psychology: Human Perception and Performance*, *21*, 344–359.
- Marian, V., Blumenfeld, H., & Boukrina, O. (2008). Sensitivity to phonological similarity within and across languages. *Journal of Psycholinguistic Research*, *37*, 141–170.
- Marslen-Wilson, W., Nix, A., & Gaskell, G. (1995). Phonological variation in lexical access: Abstractness, inference and English place assimilation. *Language and Cognitive Processes*, *10*, 285–308.
- Mayo, L. H., Florentine, M., & Buus, S. (1997). Age of second-language acquisition and perception of speech in noise. *Journal of Speech and Hearing Research*, *40*, 686–693.
- McQueen, J. M. (2005). Speech perception. In K. Lamberts & R. Goldstone (Eds.), *The handbook of cognition* (pp. 255–275). London: Sage Publications.
- McQueen, J. M., 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.
- Newman, R. S., Sawusch, J. R., & Luce, P. A. (1997). Lexical neighborhood effects in phonetic processing. *Journal of Experimental Psychology: Human Perception and Performance*, *23*, 873–889.
- Norris, D. (1994). Shortlist: A connectionist model of continuous speech recognition. *Cognition*, *52*, 189–234.
- Norris, D., Cutler, A., McQueen, J. M., & Butterfield, S. (2006). Phonological and conceptual activation in speech comprehension. *Cognitive Psychology*, *53*, 146–193.
- Norris, D., McQueen, J. M., & Cutler, A. (1995). Competition and segmentation in spoken-word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 1209–1228.
- Norris, D., McQueen, J. M., Cutler, A., & Butterfield, S. (1997). The possible-word constraint in the segmentation of continuous speech. *Cognitive Psychology*, *34*, 191–243.
- 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.
- Rüschmeyer, S.-A., Nojack, A., & Limbach, M. (2008). A mouse with a roof? Effects of phonological neighbors on processing of words in sentences in a non-native language. *Brain and Language*, *104*, 132–144.
- Salverda, A. P., Dahan, D., & McQueen, J. M. (2003). The role of prosodic boundaries in the resolution of lexical embedding in speech comprehension. *Cognition*, *90*, 51–89.

- Schouten, M. E. H. (1975). *Native-language interference in the perception of second-language vowels: An investigation of certain aspects of the acquisition of a second language*. Unpublished doctoral dissertation. Utrecht University, The Netherlands.
- Sebastián-Gallés, N., Echeverría, S., & Bosch, L. (2005). The influence of initial exposure on lexical representation: Comparing early and simultaneous bilinguals. *Journal of Memory and Language*, *52*, 240–255.
- Shatzman, K. B., & McQueen, J. M. (2006). Segment duration as a cue to word boundaries in spoken-word recognition. *Perception and Psychophysics*, *68*, 1–16.
- Soto-Faraco, S., Sebastián-Gallés, N., & Cutler, A. (2001). Segmental and suprasegmental mismatch in lexical access. *Journal of Memory and Language*, *45*, 412–432.
- Strange, W. (Ed.). (1995). *Speech perception and linguistic experience: Issues in cross-language research*. Baltimore: York Press.
- Swinney, D. A. (1979). Lexical access during sentence comprehension: (Re)consideration of context effects. *Journal of Verbal Learning and Verbal Behavior*, *18*, 645–659.
- Tabossi, P., Burani, C., & Scott, D. (1995). Word identification in fluent speech. *Journal of Memory and Language*, *34*, 440–467.
- Taft, M. (1986). Lexical access codes in visual and auditory word recognition. *Language and Cognitive Processes*, *1*, 297–308.
- Van Wijngaarden, S. J., Steeneken, H. J. M., & Houtgast, T. (2002). Quantifying the intelligibility of speech in noise for non-native listeners. *Journal of the Acoustical Society of America*, *111*, 1906–1916.
- 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.
- Weber, A., & Cutler, A. (2004). Lexical competition in non-native spoken-word recognition. *Journal of Memory and Language*, *50*, 1–25.
- Zwitserslood, P. (1989). The locus of the effects of sentential-semantic context in spoken-word processing. *Cognition*, *32*, 25–64.
- Zwitserslood, P. (1996). Form priming. *Language and Cognitive Processes*, *11*, 589–596.

## APPENDIX A

## Experimental stimuli, Experiments 1 and 3.

Experiment 1: The first column represents words and the second column near-words.

Experiment 3: The first column also represents the visual targets.

<i>Primes (Experiment 3)</i>			<i>Carrier fragments (Experiment 3)</i>		
<i>Match</i>	<i>Mismatch</i>	<i>Control</i>	<i>Match</i>	<i>Mismatch</i>	<i>Control</i>
<i>/æ/</i>					
bank	benk	skeep	superb ankle	superb enclave	glass keep
black	bleck	tring	superb lack	superb lecture	hot ring
blank	blenk	glice	superb lankiness	superb Lenkerville	big license
gram	grem	trade	snug ram	snug REM sleep	wet radar
lamp	lemp	bike	evil amplitude	evil empire	prefab icon
plank	plenk	fear	hip lankiness	hip Lenkerville	stiff ear
pram	prem	fear	deep ram	deep REM sleep	stiff ear
rank	renk	glice	clear ankle	clear enclave	big license
scratch	scretch	glice	brisk ratchbet	brisk retch	big license
slam	slem	fear	Swiss laminate	Swiss lemonade	stiff ear
smash	smesh	trade	nice mash	nice mesh	wet radar
span	spen	tring	glass pan	glass pen	hot ring
spank	spenk	bike	crisp ankle	crisp enclave	prefab icon
splash	splash	tring	nice Pleshterville <sup>a</sup>	nice Pleshterville <sup>a</sup>	hot ring
thank	think	bike	fourth ankle	fourth enclave	prefab icon
trap	trep	skeep	that rap	that rep	glass keep
<i>/ɛ/</i>					
bench	banch	sun	superb enchilada	superb anchovy	this undertone
breast	brast	trick	prefab restaurant	prefab raster	that rick
breath	brath	skeep	superb Reathly <sup>a</sup>	superb Rathly <sup>a</sup>	glass keep
chess	chass	buse	each es	each ass	arab user
chest	chast	grain	each estimate	each asteroid	big rain
death	dath	nig	good ethnic	good athlete	fun ignition
desk	dask	sun	bad escort	bad asking	this undertone
dress	drass	grain	good rescue	good raster	big rain
fresh	frash	buse	half Resbdale <sup>a</sup>	half Rasbdale <sup>a</sup>	arab user
jet	jat	trick	the hostage ate	hostage at work	that rick
press	prass	nig	sharp rescue	sharp raster	fun ignition
quest	quast	buse	dark west	dark Wast Water	arab user
smell	smal	nig	this melon	this malice	fun ignition
sweat	swatt	trick	this wetting	this Watchet	that rick
swell	swall	grain	this welder	this Walbury Hill	big rain
wealth	walth	sun	view Eltham Palace	view Althea Lake	this undertone

*Note:* In Experiment 3, some control condition primes occurred more than once; no participant, however, heard any prime more than once.

The average number of Dutch neighbours beginning with the same onset and Dutch /ɛ/ was 6.9 for /æ/ items in the prime/match column and 9.3 for /ɛ/ items,  $F(1, 31)=2.0$ , *ns*.

<sup>a</sup>Fictitious geographic names in the carrier fragments.

APPENDIX B

Experimental stimuli, Experiments 2 and 4

The first column also represents the visual targets in Experiments 2 and 4. Average number of Dutch neighbours: /æ/ 7.8, /ɛ/ 11.7,  $F(1, 11) = 1.98, ns$ .

<i>Primes (Experiment 2)</i>			<i>Carrier fragments (Experiment 2)</i>		
<i>Match</i>	<i>Mismatch</i>	<i>Control</i>	<i>Match</i>	<i>Mismatch</i>	<i>Control</i>
<i>[Not used in Experiment 4]</i>			<i>Primes (Experiment 4)</i>		
			<i>Match</i>	<i>Mismatch</i>	<i>Control</i>
/æ/					
cat	ket	pock	<i>catalogue</i>	<i>kettle</i>	<i>pocket</i>
damn	dem	al	<i>damage</i>	<i>democrat</i>	<i>albatross</i>
lamb	lem	bal	<i>laminare</i>	<i>lemon</i>	<i>balance</i>
pant	pent	synth	<i>pantomime</i>	<i>pentagram</i>	<i>synthesize</i>
tack	teck	fung	<i>tactic</i>	<i>textile</i>	<i>finger</i>
tan	ten	bul	<i>tantrum</i>	<i>tentacle</i>	<i>bulletin</i>
/ɛ/					
chess	chas	an	<i>chestnut</i>	<i>chastity</i>	<i>anarchism</i>
deaf	daf	hov	<i>definite</i>	<i>daffodil</i>	<i>hovercraft</i>
dress	dras	stupe	<i>dressage</i>	<i>drastic</i>	<i>stupid</i>
edge	adge	whis	<i>educate</i>	<i>agile</i>	<i>whistle</i>
egg	ag	lune	<i>egotist</i>	<i>agony</i>	<i>lunatic</i>
shed	shad	vir	<i>schedule</i>	<i>shadow</i>	<i>virgin</i>

Downloaded By: [Max Planck Institute for Psycholinguistic] At: 17:31 6 January 2011