Phonetic and lexical processing in a second language

ISBN: 90-76203-21-0 / 978-90-76203-21-8

Cover design: Linda van den Akker, Inge Doehring

Cover illustrations:

Cattle: reproduced by kind permission of the journal Veeteelt

Kettle: www.sportswarehouse.co.uk

Printed and bound by Ponsen & Looijen bv, Wageningen

Phonetic and lexical processing in a second language

een wetenschappelijke proeve

op het gebied van de Sociale Wetenschappen

Proefschrift

ter verkrijging van de graad van doctor

aan de Radboud Universiteit Nijmegen

op gezag van de Rector Magnificus Prof. dr. C.W.P.M. Blom,

volgens besluit van het College van Decanen

in het openbaar te verdedigen

op dinsdag 20 december 2005

des ochtends om 10:30 uur precies

door

Mirjam Elisabeth Broersma

geboren op 24 september 1976 te Groningen

Promotor:	Prof. dr. A. Cutler
Manuscriptcommissie:	Prof. dr. P.C. Muysken
	Prof. N. Sebastián-Gallés (Universitat de Barcelona)
	Dr. A.F.J. Dijkstra

The research reported in this thesis was supported by a grant from the Max-Planck-Gesellschaft zur Förderung der Wissenschaften, München, Germany, and by a Marie Curie Fellowship.

Sometimes Alex thought that if you got all the part-time mature students in the world and laid them head to toe around the line of the equator strapped down in some way so they couldn’t move, that would be a good thing.

Zadie Smith, *The autograph man*

Dankwoord

In oktober 2000 kwam ik op het MPI voor een sollicitatiegesprek en ik wist meteen dat ik wilde blijven. Ik zag mezelf al helemaal zitten, op dat instituut waar de mensen vol enthousiasme over hun onderzoek praten en waar in een heel ontspannen sfeer heel hard gewerkt wordt. Andrea, Petra en Simone namen me mee naar de kantine en vertelden me over het MPI, over de *Comprehension Group* en over het schrijven van een proefschrift, en ik wist: dit is wat ik wil. Ik kreeg de baan, ik schreef het proefschrift, en inderdaad: het was ontzettend leuk!

Mijn hartelijke dank gaat uit naar Anne, mijn promotor en begeleider. Omdat ze me aannam als *PhD student*, omdat ze mij het idee voor dit onderzoek en voor verschillende experimenten aan de hand deed, en omdat ze voor mijn eerste presentatie op een conferentie nog zenuwachtiger was dan ik. Ook wil ik haar bedanken voor het supersnelle commentaar op alle teksten die ik bij haar afleverde. Hoezo *feedback is never necessary*? Anne las mijn stukken in de auto, in de trein, in het vliegtuig, ‘s nachts, in het weekend, en één keer, denk ik, stiekem op een bruiloft.

Ten tweede wil ik James bedanken, die in al mijn experimenten prominent aanwezig was als ‘de stem’. Deelnemers aan mijn experimenten waren erg nieuwsgierig naar de spreker achter de eindeloze reeks monotone woordjes en beschreven ‘de stem’ ondermeer als een robot, een *computerized train announcer* en als een elf uit Lord of the Rings - maar vooral ook als prettig en heel duidelijk. Prettig en heel duidelijk was ook het commentaar dat James als tweede lezer op mijn stukken gaf, net als de antwoorden op de vragen waarmee ik altijd bij hem terecht kon. Ik heb er heel veel aan gehad.

Ik wil iedereen in de Comprehension Group bedanken voor de geweldige sfeer, voor het meedenken en voor alle hulp. Ik wil Roel bedanken voor zijn hulp bij verschillende onderdelen van mijn project. Petra dank ik voor het advies over fonetische categorisatie, Mirjam voor haar hulp met statistische analyses, Keren en Tau voor hun hulp met *awk*, *gawk* en de vele technische dingetjes waar ik niet uitkwam, en Rian voor alle administratieve hulp. I would like to thank Marianne for her help with running experiments.

Martijn was mijn kamergenoot gedurende het grootste deel van mijn tijd op het MPI. Samen verhuisden we van het donkerste hokje naar de grootste kamer van de afdeling, en samen sleepten we een 2 meter hoog Ghanees beeld uit de kelder om ons op onze plichten te wijzen. We waren een jaar lang samen *PhD representative* en tweemaal samen paranimf. Als het hem al de neus uitkwam heeft hij dit nooit laten blijken. Dank hiervoor!

In mijn ‘cohort’ zaten ook Marjolein, Etienne en Rachèl. Er waren lunchpauzes die een halve middag doorgingen, er waren vele etentjes, er werd gespijbeld tijdens een *summer school*, er waren twee promoties en een bruiloft, en wie weet wat er nog meer gaat komen?

Verder wil ik Anne Pier bedanken, mijn kamergenoot in de eerste maand van mijn project, toen hij mij wegwijs maakte op het instituut, en tijdens de laatste maanden, toen we beiden vochten met de laatste versies van ons proefschrift. Bijna waren we op dezelfde dag gepromoveerd, maar met twee achtereenvolgende dagen zetten we ook een nieuw instituutsrecord neer.

I would like to thank my third roommate, Anita for bringing laughter and frog colors into the office. We had cheese cake in New York, heard people sing at Starbucks, and in Brighton we spent a night together with a gigolo who made his fortune by selling scotch tape. I can’t believe the institute pays us for trips like this!

Mijn andere reisgenoten in New York, Holger en Mirjam wil ik bedanken voor de gezelligheid en de morele steun. Fijn dat die terminal pas instortte een paar uur nadat wij vertrokken waren. Dank aan Keren voor de leuke tijd in Korea en omdat zij de *ondol* nam. Ik was blij dat ik niet alleen stond met die zes regels lange vertaling van “heeft u ook vegetarisch eten?” Dank aan Claudia en Christiane voor alle goede gesprekken, zowel onder werktijd als daarbuiten.

Voor de etentjes, films, feestjes en gezelligheid wil ik alle PhDs bedanken, and I would also like to thank all the PhDs who didn’t understand the previous sentence!

Ik wil Antje Meyer bedanken voor de keren dat ik in haar lab in het Behavioral Brain Science Center aan de University of Birmingham proefpersonen mocht komen testen. I would like to thank Alan Garnham of the Laboratory of Experimental Psychology at the University of Sussex for giving me the opportunity to work in his lab for three months. It was a very pleasant and productive time.

Dank aan mijn zwager Bert, hoofdredacteur van het vakblad Veeteelt, voor de prachtige foto die op de voorkant van dit proefschrift prijkt.

Mijn broer Daniël, die twee dagen na mij zijn proefschrift verdedigt, wens ik hierbij heel veel succes.

Evelien, die mij haar auto leende om naar Frankrijk te rijden nog geen maand nadat ik mijn rijbewijs had gehaald en die het niet erg vindt als ik haar cd’s pas na een jaar teruggeef, bedankt voor het vertrouwen. Frank werd mijn vierde kamergenoot in de meest hectische fase van mijn project en bracht met zijn sereniteit en zijn luisterend oor rust. En vertrouwen en rust zijn nou net die dingen die ik nodig heb bij mijn promotie. Ik ben dan ook heel blij dat Evelien en Frank mijn paranimfen willen zijn!

En Wilfred... tja, waar zal ik eens beginnen? Ik vertel het je thuis wel.

Contents

Introduction

Chapter 1

Speech comprehension in a second language

The comprehension of speech in a second language is much more difficult than the comprehension of speech in one's native language. Whereas the comprehension of native speech is usually effortless and listeners are totally unaware of the complexity of the task they are performing, listening to a second language can make them painfully aware of this complexity.

Trying to understand speech in a second language can be so demanding that it can cause a temporary decrease in thinking ability. In a study by Takano and Noda (1993), participants performed several cognitive tasks, like calculation or tracing the way out of a printed maze. During these tasks, they were required to listen either to speech in their mother tongue or to speech in a second language which they knew well. When the participants listened to their second language, they made more errors on the cognitive tasks than when they listened to their first language. Processing the second language was so demanding that it diminished the listeners' ability to perform other cognitive tasks at the same time.

Nonnative speech comprehension is difficult in many respects. Speech in a second language may seem too fast for the listener to be able to distinguish separate words, it contains confusing speech sounds, as well as words and expressions which the listener does not know, the recognition of words takes longer than in the native language, and even when all the separate words have been recognized, it is still not always clear what the sentence means.

Some of the problems that listeners experience during the perception of a second language are easy to understand. It is obvious that second language learners, especially those who have just begun learning the language, have a smaller vocabulary than native listeners, so that they may encounter many words they do not know. Other difficulties are less transparent, but can be explained by the cognitive processes underlying speech comprehension.

Segmenting speech into separate words

For example, why is it that speech in a second language often seems faster than speech in the native language? By definition, second languages cannot always be faster than first languages, as one person's second language is another person's first language. Therefore, the impression of fastness must have a different cause. The explanation is probably that listeners do not always manage to segment speech in a second language into separate words, but perceive it as one long, indivisible stream of speech instead. Unlike in writing, where words are clearly separated by spaces, words are not separated in a comparable way in speech. Pauses in speech do not indicate word boundaries like spaces in writing: they may occur within words and they may be missing between words. Thus, listeners have to use other information to find the boundaries between the words in an utterance. Native listeners of different languages have different ways of finding those word boundaries, and the ways to do that in one's first language may not be helpful in a second language.

Listeners can use several types of strategies to facilitate the segmentation of the speech input into separate words. One set of strategies is based on the rhythmic structure of the language. These so called metrical segmentation strategies differ for different languages. In Dutch and English, most words begin with a stressed syllable. Dutch and English listeners use this information to recognize words within the speech stream (Cutler & Norris, 1988; Vroomen, Van Zon, & De Gelder, 1996). Listeners of these languages can find words more easily in the speech stream when they begin with a stressed syllable than when they begin with an unstressed syllable. On the other hand, native listeners of French and Spanish use the boundaries of syllables to find word boundaries (Cutler, Mehler, Norris, & Seguí, 1986; Sebastián-Gallés, Dupoux, Seguí, & Mehler, 1992). In Japanese, rhythmic structure is not based on syllables but on a different unit, namely the mora. For example, the brand name *Mazda* consists of three morae: ma-z-da. Japanese listeners use a metrical segmentation strategy similar to that of French and Spanish listeners, but they use mora boundaries instead of syllable boundaries to find word boundaries (Otake, Hatano, Cutler, & Mehler, 1993). Thus, there are different types of metrical segmentation strategies. However, when people listen to a second language, they do not use the strategy that is appropriate for that language, but the one they know from their first language (Cutler et al., 1986; Otake et al., 1993). This may work well in cases where the metrical structures of the first and the second language are similar, as in Dutch and English. However, in cases where the metrical structures of the two languages differ, the strategy that is appropriate for the first language is not useful for the second language. In those cases, it will be difficult for the nonnative listener to divide the speech stream into separate words. Thus, French, which has a reputation among Dutch listeners of being extremely fast, may only seem fast to Dutch listeners because Dutch metrical segmentation strategies do not work for French.

Another kind of segmentation strategy is based on the so-called phonotactic constraints of a language. Phonotactic constraints determine which speech sounds and which combinations of sounds can occur where. For example, in English, syllables can start with /sl/ (e.g., *sleep*), but they cannot start with /nl/. Therefore, if English listeners hear a combination of the sounds /nl/, they could infer that these sounds were part of different syllables (e.g., *unless*) and possibly of different words (e.g., *on loan*). Indeed, English listeners have been shown to use this kind of information to find words in a speech stream (Weber, 2001). They found words starting with a /l/ more easily when they were preceded by a /n/ than when they were preceded by a /s/. Listening to their second language English, German listeners used the phonotactic constraints of English as well. However, they also used phonotactic constraints which were specific for German, and which were not helpful in English (Weber, 2001). This interference of phonotactic constraints from the native language may complicate the segmentation of speech in a second language into separate words.

Recognizing speech sounds

Another problem for the nonnative listener is that the speech sounds of a second language can be very difficult to distinguish. A notorious example of confusable sounds is that of /r/ and /l/ for Japanese listeners. These sounds are not only difficult to pronounce for Japanese listeners (Flege, Takagi, & Mann, 1995), but it is also difficult for them to hear the difference (Best & Strange, 1992). This is because in Japanese there is no /r/ or /l/, but there is a sound somewhere between /r/ and /l/.

Interestingly, recognizing speech sounds from other languages is something infants are often better at than adults. Infants can initially distinguish between any pair of speech sounds very well, and even between sounds from languages they have never heard before. However, when they learn more about the sounds of their own language, they lose their sensitivity to the sounds of other languages. For example, English infants could hear the difference between two Hindi consonants when they were six to eight months old, but not anymore when they were eleven to thirteen months old (Werker & Lalonde, 1988). They could hear the difference between German vowels when they were four months old, but not anymore when they were six months old (Polka & Werker, 1994).

This does not imply that, once a listener has grown up, all the speech sounds of a second language are difficult to distinguish. For adults, some sounds of a second language may still be easy to distinguish. For example, it is often easy to hear the difference between two sounds that are similar to sounds in the native language (Best & Strange, 1992). Alternatively, speech sounds in a second language may be so different from the sounds of the native language that they are not even perceived as speech. As those sounds are not confused with the speech sounds of the native language, they are also easy to distinguish. Thus, English listeners could distinguish Zulu click sounds very well (Best, McRoberts, & Sithole, 1988). However, other speech sounds are very difficult for nonnative listeners. It is difficult to hear the difference between two sounds if the native language has only one similar sound. If one of them sounds more like the native language sound than the other does, this makes it slightly easier for the listener to distinguish them (Best & Strange, 1992). The most difficult situation arises when the two sounds from the second language strongly resemble one sound from the native language, as for /r/ and /l/ in Japanese. In that case, it is extremely difficult for the nonnative listener to tell them apart (Best & Strange, 1992).

Recognizing words

Further, the recognition of words is more difficult in a second language than in the native language. This is of course the case for words that the listener does not know well, but also for other words. To make this clear, a short explanation is needed about how words are recognized.

All the words that a listener knows are stored in the brain, in the so-called mental lexicon. When a person listens to speech, words in the mental lexicon are activated. Apart from the words that the speaker intended (e.g., the Dutch word *kapitein*), many other similar words (e.g., *kapitaal*) are activated as well (Zwitserlood, 1989). While these words in the mental lexicon are active, they compete for recognition with the other active words, until one of them is recognized. Normally, the word that is recognized is the word the speaker really used. Although it may sound laborious, this is an efficient way to recognize spoken words. However, it has one drawback. The higher the number of words that are activated, the harder it is to recognize the intended word (McQueen, Norris, & Cutler, 1994; Norris, McQueen, & Cutler, 1995).

In the comprehension of a second language, not only words from the second language, but also words from the listener's first language can be activated. This has been shown in several studies. When Dutch listeners listened to English, the word *leaf* activated not only the word *leaf* in the mental lexicon, but also the similar sounding Dutch word *lief* (Schulpen, Dijkstra, Schriefers, & Hasper, 2003). When Dutch listeners heard the English word *desk*, the Dutch word *deksel*, which begins in a similar way, was also activated (Weber & Cutler, 2004). Similarly, for listeners whose first language was Russian, hearing the English word *marker* led to the activation of the Russian word *marku* (Marian, Spivey, & Hirsch, 2003). Thus, when people listen to a second language, words from their first language are also activated. As it is harder to recognize spoken words when more other words are activated, this makes the recognition of words in a second language more difficult.

Understanding sentences

A final hurdle in nonnative listening is to understand the meaning of the utterance as a whole. In order to understand the meaning of a sentence, listeners must recognize at least part of the words in the sentence. However, so-called prosodic information like sentence accent also provides information about its meaning.

For example, the following sentences (from Akker & Cutler, 2003), which only differ in the sentence accent, imply a different meaning: (1) *The tourist DIDn't fly home.* (2) *The tourist didn't FLY home.* English and Dutch listeners were found to use prosodic information in similar ways when they listened to their native language. However, Dutch listeners processed prosodic information less efficiently when listening to their second language English (Akker & Cutler, 2003). A less efficient use of prosodic information may make it more difficult to understand the meaning of sentences in a second language than in the native language.

Prosodic information can also be used to determine whether a sentence was meant literally or figuratively. Some sentences, like *it broke the ice* or *the coast is clear* can be interpreted literally or as an idiomatic expression. Native listeners can tell from the way in which a sentence is pronounced whether the speaker intended the literal meaning or the figurative meaning (Vanlancker-Sidtis, 2003). However, listeners cannot do this so well when they are listening to a second language. Even listeners who were very proficient in their second language English were not good at choosing between the literal and the figurative meaning of a sentence, and listeners who were less proficient in English could not do it at all (Vanlancker-Sidtis, 2003). Therefore, in a second language, listeners may have to use other information, like the conversational context, to choose between the literal and the figurative meaning of a sentence.

This dissertation

This dissertation further examines two of the aspects of speech comprehension that may be difficult in a second language, namely the recognition of speech sounds and the recognition of words. The steps that lead to the recognition of speech sounds are called phonetic processing. The steps that lead to the recognition of words are called lexical processing. Hence the title of this dissertation: ‘Phonetic and lexical processing in a second language’.

First, the recognition of two English speech sounds that were expected to be very difficult to distinguish for Dutch listeners was investigated. These speech sounds are the vowels in the words *lamp* and *desk*, which are phonetically written as /æ/ and /ɛ/. Like /r/ and /l/ in Japanese, Dutch does not have the same /æ/ and /ɛ/ sounds that English has, but one sound that is somewhere in between, namely the vowel in the Dutch word *pet*. Therefore, the expectation was that Dutch listeners would find it difficult to tell the two English sounds apart. It was investigated how well Dutch listeners could distinguish between the two English speech sounds, and whether they found this more difficult than listeners whose native language was English.

It was also investigated how well Dutch listeners could distinguish between the English sounds /z/ and /s/, /v/ and /f/, /b/ and /p/, and /d/ and /t/. These sounds are quite similar in English and Dutch, and therefore they might be easy for Dutch listeners to distinguish. However, in Dutch, at the end of a word, /z/, /v/, /b/, and /d/ are always pronounced as /s/, /f/, /p/, and /t/. For example, *honden* is pronounced with a /d/, but *hond* is pronounced with a /t/. Therefore, in Dutch, listeners never have to distinguish between those four sound pairs at the end of a word. In English on the other hand, all those sounds can occur at the end of a word. For example, *robe* is pronounced with a /b/ and *rope* with a /p/. It might be difficult for Dutch listeners to distinguish between these sounds at the end of words in English. Therefore, it was investigated whether Dutch listeners found it more difficult to distinguish between those English sounds at the end of words than at the beginning of words, and whether they found this more difficult than English listeners did.

Next, the recognition of words was investigated. As described above, it is harder to recognize spoken words when more other words are activated in the mental lexicon. When people listen to a second language, words from their first language are also activated, which makes the recognition of words in a second language more difficult. However, it is also possible that when people listen to a second language, more words from that second language are activated for them than for native listeners of that language. When a Dutch listener and an English listener hear the same English speech, more English words may be activated for the Dutch listener than for the English listener. This could also make the recognition of words in a second language more difficult. This additional activation of English words for Dutch listeners could occur in several situations:

If a Dutch listener heard somebody say *daf*, he could easily think that he had heard the word *deaf*. Similarly, *lemp* and *glite* could be interpreted as *lamp* and *glide*. Items like *daf*, *lemp*, and *glite*, which are very similar to real words, are called near-words in this dissertation. Of course, English speakers never say *daf*, as this is not a correct word, but they may say *DAFfodil*, which starts with the near-word *daf*. They may talk about the *eviL EMPire*, which contains the near-word *lemp*, or about a *biG LIGHT*, which contains the near-word *glite*. When a listener hears a near-word like *daf*, a word like *deaf* may be activated in the mental lexicon. This may happen more for Dutch listeners than for English listeners, for example because Dutch listeners may not be able to hear the difference between the near-word and the word (for *daf* and *deaf*), or because they may not pay attention to the difference (for *glite* and *glide*). Therefore, it was investigated whether near-words caused more activation of words in the mental lexicon for Dutch listeners than for English listeners.

Further, the recognition of words which are almost the same or which begin almost the same was investigated. For example, the first parts of the words *daffodil* and *deficit* sound almost the same. Hearing the first part of *daffodil* may lead to the activation of *deficit* in the mental lexicon, and this might happen more for Dutch listeners than for English listeners, as Dutch listeners may not hear the difference. The words *flash* and *flesh*, or *robe* and *rope* are even more similar, and it is even possible that Dutch listeners could not distinguish between them at all. It was investigated whether hearing (the first part of) a word caused more activation of the similar word in the mental lexicon for Dutch listeners than for English listeners. If this was the case, this would mean that often more English words would be activated for Dutch listeners than for English listeners when they were listening to English speech. This would be yet another process that could make the recognition of words in a second language more difficult.

References

Akker, E., & Cutler, A. (2003). Prosodic cues to semantic structure in native and nonnative listening. *Bilingualism: Language and Cognition, 6*, 81-96.

Best, C. T., McRoberts, G. W., & Sithole, N. M. (1988). Examination of perceptual reorganization for nonnative speech contrasts: Zulu click discrimination by English-speaking adults and infants. *Journal of Experimental Psychology: Human Perception and Performance, 14*, 345-360.

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

Cutler, A., Mehler, J., Norris, D., & Seguí, J. (1986). The syllable's differing role in the segmentation of French and English. *Journal of Memory and Language, 25*, 385-400.

Cutler, A., & Norris, D. (1988). The role of strong syllables in segmentation for lexical access. *Journal of Experimental Psychology: Human Perception and Performance, 14*, 113-121.

Flege, J. E., Takagi, N., & Mann, V. (1995). Japanese adults can learn to produce English /ɹ/ and /l/ accurately. *Language and Speech, 38*, 25-55.

Marian, V., Spivey, M., & Hirsch, J. (2003). Shared and separate systems in bilingual language processing: Converging evidence from eyetracking and brain imaging. *Brain and Language, 86*, 70-82.

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.

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.

Otake, T., Hatano, G., Cutler, A., & Mehler, J. (1993). Mora or syllable? Speech segmentation in Japanese. *Journal of Memory and Language, 32*, 258-278.

Polka, L., & Werker, J. F. (1994). Developmental changes in perception of nonnative vowel contrasts. *Journal of Experimental Psychology: Human Perception and Performance, 20*, 421-435.

Schulpen, B., Dijkstra, T., Schriefers, H. J., & Hasper, M. (2003). Recognition of interlingual homophones in bilingual auditory word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 29*, 1155-1178.

Sebastián-Gallés, N., Dupoux, E., Seguí, J., & Mehler, J. (1992). Contrasting syllabic effects in Catalan and Spanish. *Journal of Memory and Language, 31*, 18-32.

Takano, Y., & Noda, A. (1993). A temporary decline of thinking ability during foreign language processing. *Journal of Cross-Cultural Psychology, 24*, 445-462.

Vanlancker-Sidtis, D. (2003). Auditory recognition of idioms by native and nonnative speakers of English: It takes one to know one. *Applied Psycholinguistics, 24*, 45-57.

Vroomen, J., Van Zon, M., & De Gelder, B. (1996). Cues to speech segmentation: Evidence from juncture misperception and word spotting. *Memory and Cognition, 24*, 744-755.

Weber, A. (2001). *Language-specific listening: The case of phonetic sequences.* Unpublished doctoral dissertation, Nijmegen University, The Netherlands.

Weber, A., & Cutler, A. (2004). Lexical competition in non-native spoken-word recognition. *Journal of Memory and Language, 50*, 1-25.

Werker, J. F., & Lalonde, C. E. (1988). Cross-language speech perception: Initial capabilities and developmental change. *Developmental Psychology, 24*, 672-683.

Perception of familiar contrasts in unfamiliar positions

Chapter 2

Broersma, M. (2005). Perception of familiar contrasts in unfamiliar positions. Journal of the Acoustical Society of America, 117, 3890-3901.

Abstract

This paper investigates the perception of nonnative phoneme contrasts which exist in the native language, but not in the position tested. Like English, Dutch contrasts voiced and voiceless obstruents. Unlike English, Dutch allows only voiceless obstruents in word-final position. Dutch and English listeners' accuracy on English final voicing contrasts and their use of preceding vowel duration as a voicing cue were tested. The phonetic structure of Dutch should provide the necessary experience for a native-like use of this cue. Experiment 1 showed that Dutch listeners categorized English final /z/-/s/, /v/-/f/, /b/-/p/, and /d/-/t/ contrasts in nonwords as accurately as initial contrasts, and as accurately as English listeners did, even when release bursts were removed. In Experiment 2, English listeners used vowel duration as a cue for one final contrast, although it was uninformative and sometimes mismatched other voicing characteristics, whereas Dutch listeners did not. Although it should be relatively easy for them, Dutch listeners did not use vowel duration. Nevertheless, they attained native-like accuracy, and sometimes even outperformed the native listeners who were liable to be misled by uninformative vowel duration information. Thus, native-like use of cues for nonnative but familiar contrasts in unfamiliar positions may hardly ever be attained.

Introduction

In 1939, Trubetzkoy (reprinted as Trubetzkoy, 1977) observed that the sounds of a foreign language often get misinterpreted, because they go through the “phonological sieve” of the native language. Later research has proven Trubetzkoy right. The Perceptual Assimilation Model (PAM) (Best, 1994; Best, McRoberts, & Sithole, 1988) describes how listeners assimilate nonnative speech sounds to the native category that is perceptually most similar. The PAM predicts which nonnative speech sounds will be difficult to distinguish, based on the similarities and dissimilarities of the phonological structures of the native and the nonnative language. The most difficult distinction is that between nonnative speech sounds which match a single native category equally well. If the nonnative language has two categories where the native language has only one in the same phonetic space, both nonnative speech sounds will be assimilated to a single category. This is the case, for example, with Japanese listeners’ perception of English /r/ and /l/ (Best & Strange, 1992). According to the PAM the easiest distinction is that between nonnative speech sounds which are assimilated to two separate native categories. As the nonnative contrast corresponds to a native contrast, it is easy to perceive.

However, languages not only have a phoneme inventory, they also have their own language-specific phonotactic constraints. The perception of nonnative contrasts not only depends on the presence or absence of similar speech sounds in the native language, but also on native-language phonotactic constraints. This was demonstrated in a study of Chinese listeners’ perception of the English /d/-/t/ contrast in word-final position (Flege, 1989). Chinese has a /d/-/t/ contrast, but not in word-final position. Word-initial /d/ and /t/ are not distinguished by closure voicing in Chinese, but on the basis of information in the release burst (Flege, 1989). Flege (1989) found that Chinese learners of English categorized unedited tokens of English word-final /d/ and /t/ almost as accurately as the native English listeners did. Their performance hardly decreased when closure voicing was removed, but was strongly affected by removal of the release burst. Flege concluded that the Chinese listeners used Chinese word-initial cues to distinguish between English /d/ and /t/ in word-final position. Flege and Wang (1989) showed that not only experience with the contrast itself, but native-language experience with any word-final stops influenced the perception of the word-final stop voicing contrast. Neither Cantonese Chinese nor Mandarin Chinese has a word-final stop voicing contrast, but /p,t,k/ can occur word-finally in Cantonese, whereas Mandarin does not permit any word-final obstruents. Flege and Wang (1989) found that native listeners of Cantonese distinguished the English final /d/-/t/ contrast more accurately than native listeners of Mandarin did, which they attributed to the Mandarin listeners’ lack of native language experience with word-final obstruents.

Further, the perception of nonnative contrasts may depend not only on the presence or absence of similar phonemes in the native language, but also on the presence of utterly different contrasts. Crowther and Mann (1992) showed that the use of perceptual cues for a particular nonnative contrast may depend on the use of the same cues for other phoneme distinctions in the native language. Like Mandarin Chinese, Japanese has a /d/-/t/ contrast, and does not permit word-final stops. Whereas Japanese has long and short vowels, Mandarin Chinese does not have this distinction. Crowther and Mann tested the perception of the English word-final /d/-/t/ contrast by Japanese and Mandarin learners of English. The Japanese listeners showed a greater sensitivity to the duration of the vowel preceding the final consonant and categorized the English final /d/-/t/ contrast more accurately than the Mandarin listeners did.

Thus, the categorization of a nonnative contrast which exists in the native language, but in a position where it does not occur in the native language, seems to benefit from native-language experience with one of the phonemes of the contrast in the relevant position and from experience with relevant perceptual cues. These findings suggest that the potential for accurate and native-like categorization of a nonnative but familiar contrast in an unfamiliar position is highest for native listeners of a language which provides such experience. Of all languages that contrast voiced and voiceless obstruents but not in word-final position, those languages which allow for either voiced or voiceless obstruents in word-final position, and in which vowel duration is used as a cue (for any phoneme contrast), offer the best preparation for accurate categorization of the word-final obstruent voicing contrast and for the use of vowel duration as a cue. As Dutch has a distinction between voiced and voiceless obstruents in word-initial and -medial position, allows for voiceless obstruents in word-final position, and also provides experience with the use of vowel duration as a cue for several phoneme distinctions, native listeners of Dutch should be well prepared to learn to distinguish English voiced and voiceless word-final obstruents as a familiar contrast in an unfamiliar position, and to use vowel duration as a cue. Especially advanced learners of English can be expected to have learned to do this, through combining their native and nonnative language experience. Therefore, this paper investigates whether Dutch listeners with a high level of proficiency in English categorize English final obstruent voicing contrasts with a native-like level of accuracy and with a native-like use of the vowel duration cue. It provides a test of the perception of a nonnative but familiar contrast in an unfamiliar position by listeners with a language background that is most suitable for the task.

Dutch and English share four pairs of voiced and voiceless obstruents: the alveolar and labiodental fricatives /z/, /s/, /v/, and /f/, and the bilabial and alveolar stops /b/, /p/, /d/, and /t/. Unlike English, Dutch neutralizes voicing distinctions in syllable-final, prepausal position (Booij, 1995). Thus, although in Dutch obstruent voicing is a relevant contrast in word-initial and -medial position, Dutch has no word-final voicing contrasts. Dutch does allow for /s,f,p,t/

in word-final position. Further, Dutch distinguishes between long and short vowels (Booij, 1995). As part of the difference between long and short vowels is phonetic vowel duration, Dutch listeners are familiar with the assessment of this cue. Dutch listeners even have native-language experience with the use of vowel duration as a cue to word-medial obstruent voicing. In Dutch, vowels preceding a medial voiced consonant are slightly longer than vowels preceding a medial voiceless consonant. According to Slis and Cohen (1969a) the average difference is 30 ms before stops and 40 ms before fricatives. Van den Berg (1989) found that Dutch listeners used vowel duration to decide on the voicing of intervocalic two-obstruent sequences, although it was not among the most important perceptual cues. A study by Jongman, Sereno, Raaijmakers, and Lahiri (1992) suggests that Dutch listeners may be able to generalize their knowledge about the relationship between vowel duration and word-medial obstruent voicing to the case of word-final obstruents. In this study, Dutch listeners categorized vowels from a vowel length continuum as long or short. Stimuli corresponded to the Dutch words /stad/ and /sta:t/, and /zat/ and /za:d/, in which vowel length and underlying voicing of the final consonant are crossed. The surface word-final consonant was always voiceless. The location of the phoneme boundary differed between the two continua, suggesting that the perception of ambiguous vowel duration depended on the underlying voicing of the word-final stop.

In English, the difference in vowel duration before voiced and voiceless obstruents is larger than in Dutch. Peterson and Lehiste (1960) found a difference of 96 ms before word-final stops and 148 ms before word-final fricatives. There is extensive evidence for the great importance of preceding vowel duration for the perception of voicing of word-final obstruents in English (e.g., Raphael, 1972). Although the role of vowel duration as a cue to voicing seems to be smaller in Dutch than in English, Dutch listeners' familiarity with the cue in word-medial position may facilitate its use in word-final position in English.

As their native language has not provided them with any knowledge about the relevant acoustic cues for voicing in final position, Dutch listeners may try to identify the voicing of English final obstruents with the aid of the perceptual cues they rely on for Dutch initial and intervocalic voicing contrasts. This may be quite successful, as Dutch and English obstruents have a high degree of articulatory similarity, and the perceptual cues that signal the voicing distinctions overlap to some extent. Van Alphen and Smits (2004) showed that Voice Onset Time (VOT), specifically the presence or absence of prevoicing, is the strongest cue to initial stop voicing in Dutch. In the absence of prevoicing, voicing judgements for labials relied most strongly on the extent of F0 change into the following vowel, and for alveolars on the spectral center of gravity of the burst. Other significant cues were the duration and power of the burst. For intervocalic obstruents, presence or absence of vocal fold vibration (Slis & Cohen, 1969b; Slis & Van Heugten, 1989), closure duration for stops (Kuijpers, 1996; Slis & Cohen, 1969a) and frication duration for fricatives (Slis & Van Heugten, 1989) have been shown to influence the perception of voicing. For intervocalic two-obstruent sequences, presence or absence of vocal fold vibration during the closure of the two obstruents is the most important cue (Van den Berg, 1989). Closure duration of the second consonant, duration of the preceding vowel, and for fricatives the intensity of frication noise play a smaller but significant role (Van den Berg, 1989).

All of the above-mentioned cues have been found to be used by English listeners to distinguish voiced and voiceless obstruents in word-final position (for a review, see Watson, 1983), and Dutch listeners may use their knowledge about Dutch voiced and voiceless obstruents to make the same distinction for final English obstruents. However, there are differences between Dutch and English obstruents, and thus between the critical values of the cues, and the weight attributed to each cue for optimal identification.

For example, the two languages differ in the critical value of VOT for voicing of initial stops. English contrasts voiceless unaspirated and voiceless aspirated stops (Lisker & Abramson, 1964), and the duration of the voicing lag is a cue to voicing in English (e.g., Watson, 1983). Van Alphen and Smits (2004) found that 75 % of Dutch voiced initial stops were produced with a voicing lead, and that the presence or absence of prevoicing was the strongest perceptual cue for initial stop consonant voicing in Dutch. Although initial stops without prevoicing were not automatically categorized as voiceless, but were assessed on the basis of other cues (as described above), initial stops without prevoicing were misperceived more often than prevoiced stops (37 % vs. 1 %). In English, initial voiced stops are less often prevoiced than in Dutch. Smith (1978) found that bilabial voiced stops were prevoiced 56 %, and alveolar stops 50 % of the time in careful speech. Therefore, Dutch listeners may misperceive English initial voiced stops relatively often.

Another difference between Dutch and English is the importance of the duration of the preceding vowel as a cue to obstruent voicing. If Dutch listeners process English final obstruents in the same way they process Dutch obstruents, they may not attribute as much weight to vowel duration as English listeners do. This may not be a problem when enough other cues are available, but it may lead to less accurate categorization of unreleased stops. In English, final stops are often produced without a release burst (Byrd, 1993). English listeners have little difficulty identifying the voicing of stops without a release burst (e.g., Flege & Hillenbrand, 1987), which may be explained by the redundancy of information in the speech signal. However, if Dutch listeners use vowel duration as a voicing cue less than English listeners do, the Dutch listeners may have more difficulty identifying English final obstruents without a release burst.

In Experiment 1, Dutch and English listeners' categorization of the British English obstruent voicing contrasts /z/-/s/, /v/-/f/, /b/-/p/, and /d/-/t/ was investigated in initial and final position in nonwords. It was investigated whether Dutch listeners had a preference for identifying English final obstruents as voiceless, as Dutch allows voiceless but not voiced obstruents in word-final position. The effect of removal of the release burst was investigated for the final stops. For reasons of comparison, a contrast which was expected to be difficult to distinguish for Dutch listeners was included in the experiment, namely the English /æ/-/ɛ/ contrast. The PAM predicts that this phoneme pair belongs to the set of most difficult English contrasts for Dutch listeners. Standard southern British English distinguishes two open mid-front unrounded vowels, whereas Dutch has only one vowel in this part of the vowel space. Although the Dutch vowel is denoted as /ɛ/, it is lower than the English /ɛ/, so that it is located between English /ɛ/ and /æ/. As Dutch listeners will assimilate both English vowels to the single Dutch category, the distinction between the phonemes is expected to be difficult. Indeed, British English /æ/ and /ɛ/ have been found to be difficult to distinguish for Dutch listeners (Schouten, 1975).

Dutch and English listeners' use of vowel duration as a cue to final obstruent voicing was further investigated in Experiment 2. If Dutch listeners use vowel duration as a cue less than English listeners do, Dutch listeners may find it easier to ignore vowel duration when this cue is made unreliable than English listeners do. In Experiment 2 it was investigated whether Dutch and English listeners relied on vowel duration as a cue to final obstruent voicing when this cue was uninformative and when it mismatched with other information in the signal.

Experiment 1

Method

Participants

Twenty native speakers of Dutch and 20 native speakers of British English took part in the experiment. The Dutch participants had a high level of proficiency in English as a second language. They had received on average 7 years of English instruction in primary and secondary education. The English participants did not know any Dutch. 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. None reported any hearing loss. All were volunteers and received a small fee for participation.

Materials

The vowel contrast /æ/-/ɛ/ was to be tested in one position and the four consonant contrasts /z/-/s/, /v/-/f/, /b/-/p/, and /d/-/t/ in two positions. Therefore, nine pairs of monosyllabic CVC items were selected. Each pair differed in one phoneme pair, corresponding to the contrast to be tested. The non-target consonants in the CVC items were obstruents, in order to minimize their influence on the target sounds. All items were nonwords in Dutch and English, according to the CELEX database (Baayen, Piepenbrock, & Gulikers, 1995). The items are presented in Table 1.

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 soundproof booth with a Sennheiser microphone and stored directly onto a computer at a sample rate of 16 kHz. For each target phoneme in each position, three tokens were extracted from the file with the speech editor Praat. For the items with a final stop, only tokens with a clearly audible release burst were selected. These tokens were kept unedited for the condition with release burst, and for the condition without release burst the signal was truncated at the last positive zero crossing before the release burst.

Acoustic measurements were made of several characteristics which may be relevant for the distinction of the contrasts. The results are presented in Table 2 for the target vowels, in Table 3 for the fricatives and in Table 4 for the stops.

Table 1. Experiment 1 items.

/æ/-/ɛ/	/f**æ**f/ - /f**ɛ**f/
Initial /z/-/s/	/**z**i:f/ - /**s**i:f/
Final /z/-/s/	/fu:**z**/ - /fu:**s**/
Initial /v/-/f/	/**v**u:k/ - /**f**u:k/
Final /v/-/f/	/ku:**v**/ - /ku:**f**/
Initial /b/-/p/	/**b**o:f/ - /**p**o:f/
Final /b/-/p/	/fi:**b**/ - /fi:**p**/
Initial /d/-/t/	/**d**i:s/ - /**t**i:s/
Final /d/-/t/	/fo:**d**/ - /fo:**t**/

Table 2. Experiment 1, acoustic measures of stimuli with /æ/ and /ɛ/: Mean F1 steady state frequency (Hz), mean F2 steady state frequency (Hz), and mean vowel duration (ms).

	/æ/	**/ɛ/**
F1	824	744
F2	1602	1823
Vowel duration	167.4	131.6

Table 3. Experiment 1, acoustic measures of stimuli with initial and final fricatives: Mean vowel duration (ms), mean fricative duration (ms), and mean fricative power above 500 Hz (logarithm of the spectral power of the frication noise above 500 Hz in Pa^2).

	Initial				**Final**			
	/z/	**/s/**	**/v/**	**/f/**	**/z/**	**/s/**	**/v/**	**/f/**
Vowel duration	-	-	-	-	258.7	130.5	264.9	118.5
Fricative duration	112.8	178.3	111.2	160.9	160.5	257.8	144.8	232.7
Fricative power	-2.2	-2.2	-3.1	-3.1	-2.6	-2.2	-3.6	-3.3

Table 4. Experiment 1, acoustic measures of stimuli with initial and final stops: Proportion of initial stops with prevoicing, mean vowel duration (ms), mean F1 offset frequency (Hz), mean closure duration (ms), proportion of final stops with voicing during closure, mean closure voicing duration (as a percentage of total closure duration), and mean burst duration (ms).

	Initial				**Final**			
	/b/	**/p/**	**/d/**	**/t/**	**/b/**	**/p/**	**/d/**	**/t/**
Prevoicing	2/3	0	3/3	0	-	-	-	-
Vowel duration	-	-	-	-	185.7	112.7	252.8	133.3
F1 offset frequency	-	-	-	-	246	261	272	310
Closure duration	-	-	-	-	82.1	117.8	59.1	113.6
Closure voicing	-	-	-	-	3/3	0	3/3	3/3
Closure voicing duration (%)	-	-	-	-	94.2	0	100	27.5
Burst duration	10.8	22.6	16.3	18.7	67.5	72.0	72.0	122.9

Design

Each fricative contrast occurred in initial and final position. The stops occurred in three conditions: initial position, final position with release burst, and final position without release burst. The order of presentation of the initial and final positions was counterbalanced. As the items in the two final conditions were based on the same tokens, the final with release burst condition always occurred after the final without release burst condition. The target phonemes /s/ and /f/ also occurred as non-targets in stimuli for other contrasts. The blocks were ordered such that the subjects were not exposed to a phoneme before the contrast it was part of was being tested.

The items were presented in 11 blocks, each block representing one phoneme contrast in one condition. Each block consisted of four repetitions of six tokens, semi-randomized such that the same phoneme occurred maximally five times in succession and the same token maximally once.

Procedure

Participants were tested one at a time in a quiet room. They were informed in their native language that they would hear a series of non-words, which would be similar except for one sound. They were instructed to decide which one of two alternatives this sound was, and to indicate their response with a button press. Before each block, they received further information about the two response alternatives in that block, and about the position of the target phoneme. They were not instructed about the truncation in the condition without release burst. Before the /æ/-/ɛ/ block, participants heard some examples of non-words containing these phonemes, to make it clear, particularly to the Dutch participants, which sounds were intended. The other phonemes were not expected to cause uncertainty, and were not illustrated with examples. Each block started with six practice trials. The response buttons were labeled “A” and “E”, “Z” and “S”, “V” and “F”, “B” and “P”, or “D” and “T”, respectively. The experiment was controlled with NESU (Nijmegen Experiment Set-Up) experimental software. Stimuli were presented binaurally over Sennheiser closed headphones at a comfortable listening level, one at a time. Participants responded by pressing one of two response buttons. No time limit was imposed for the responses. After each button press, presentation of the next item started.

Results and discussion

One response with a reaction time (RT) longer than 10,000 ms due to a technical error was removed. One Dutch subject gave only "v" responses for the /v/-/f/ contrast in final position. All responses of this subject on both /v/-/f/ contrasts were removed from the analysis. Mean percentages of correct responses are presented in Table 5. The sensitivity measure *d'* was calculated for each subject, for each contrast, and each condition separately, with a correction for near-perfect sensitivity (MacMillan & Creelman, 1991). Next, log β was calculated to investigate possible biases (McNicol, 1972). Mean values of *d'* and log β are presented in Table 6.

For the /æ/-/ɛ/ contrast, an analysis of variance (ANOVA) showed that the *d'*s of the English listeners were significantly larger than those of the Dutch listeners ($F\ (1, 39) = 7.59$, $p < .01$), indicating a higher sensitivity of the English listeners. However, a t-test showed that the Dutch listeners performed amply above chance (*d'* = 0), with *d'*s significantly larger than 3 ($t\ (19) = 5.38$, $p < .001$). There was no effect of native language on bias ($F\ (1, 39) = 2.21$, $p > .1$).

For the /z/-/s/ contrast, no interaction between condition and native language ($F\ (1, 38) = 1.93$, $p > .1$), and no main effects of condition ($F\ (1, 38) < 1$) and native language ($F\ (1, 38) < 1$) were found. No difference in bias between the language groups was found for initial position ($F\ (1, 39) < 1$) or for final position ($F\ (1, 39) < 1$).

Table 5. Experiment 1 results: Mean percentage of correct responses as a function of participants' native language and condition.

	Dutch				English			
	Medial	**Initial**	**Final released**	**Final dereleased**	**Medial**	**Initial**	**Final released**	**Final dereleased**
/æ/	96	-	-	-	97	-	-	-
/ɛ/	94	-	-	-	100	-	-	-
/z/	-	96	96	-	-	97	93	-
/s/	-	96	98	-	-	98	95	-
/v/	-	95	94	-	-	98	97	-
/f/	-	92	100	-	-	98	99	-
/b/	-	86	96	92	-	99	98	96
/p/	-	98	98	93	-	99	98	96
/d/	-	99	94	95	-	97	96	97
/t/	-	99	98	94	-	98	95	93

Table 6. Experiment 1 results: Mean d' *and log* β *as a function of participants' native language and condition. (Higher values of* d' *indicate higher sensitivity. Negative values of log* β *indicate a bias towards the first, and positive values towards the second phoneme of a contrast.)*

	Dutch		English	
	d'	**log *β***	***d'***	**log *β***
/æ/-/ɛ/	4.21	-0.28	4.96	0.59
Initial /z/-/s/	4.48	-0.25	4.57	0.00
Final /z/-/s/	4.55	0.33	4.17	-0.07
Initial /v/-/f/	4.20	-0.64	4.83	0.00
Final /v/-/f/	4.61	1.23	4.79	0.43
Initial /b/-/p/	3.93	1.76	5.14	0.00
Final /b/-/p/ released	4.68	0.34	4.72	0.02
Final /b/-/p/ dereleased	3.85	0.42	4.31	0.11
Initial /d/-/t/	5.11	-0.02	4.73	0.42
Final /d/-/t/ released	4.45	0.39	4.21	-0.31
Final /d/-/t/ dereleased	3.90	0.09	4.11	-1.00

For the /v/-/f/ contrast, no interaction between condition and native language (F (1, 37) = 1.22, $p > .1$), and no main effects of condition (F (1, 37) < 1) and native language (F (1, 37) = 2.40, $p > .1$) were found. No difference in bias between the language groups was found for initial position (F (1, 39) = 1.07, $p > .1$) or for final position (F (1, 38) = 2.27, $p > .1$).

For the /b/-/p/ contrast, a significant interaction between condition and native language was found (F (2, 76) = 4.45, $p < .05$). A planned comparison of initial position and final position with release burst yielded a significant interaction between condition and native language (F (1, 38) = 8.65, $p < .01$). Therefore, separate analyses were performed for both conditions and both language groups. In initial position, the English listeners' sensitivity was significantly higher than the Dutch listeners' sensitivity (F (1, 39) = 19.75, $p < .001$). In final position with release burst, there was no effect of native language (F (1, 39) < 1). Comparing initial position and final position with release burst for the Dutch listeners only, a significantly lower sensitivity was found for initial position (F (1, 19) = 5.68, $p < .05$). For the English listeners, there was no difference between initial position and final position with release burst (F (1, 19) = 2.98, $p > .1$).

In a planned comparison of final position with release burst and final position without release burst, no interaction was found between condition and native language (F (1, 38) = 1.21, $p > .1$). The effect of condition was significant (F (1, 38) = 10.69, $p < .01$), with *d'* being larger for final position with release burst than for final position without release burst. There was no significant effect of native language (F (1, 38) < 1).

For initial position, the effect of native language on bias was significant (F (1, 39) = 13.94, $p < .001$), with a bias towards "p" responses for the Dutch listeners, and no bias (log $\beta = 0$) for the English listeners. Neither the analysis of log β for final position with release burst (F (1, 39) < 1) nor that without release burst (F (1, 39) < 1) yielded a significant effect.

The Dutch listeners' bias towards "p" responses in initial position suggested that their low sensitivity for the initial contrast resulted from a high number of errors on /b/ items rather than on /p/ items. As Table 5 shows, the Dutch listeners' percentage of correct responses for /p/ items was similar to that of the English listeners (98 % vs. 99 %), whereas the Dutch listeners' percentage of correct responses for /b/ items was only 86 %, compared to 99 % for the English listeners. Acoustical examination of the /b/ items showed that two of the tokens were produced with prevoicing, and the third without prevoicing (Table 4). The tokens with prevoicing received 99 % and 96 % correct responses from the Dutch listeners, whereas the token without prevoicing received only 63 % correct responses from the Dutch listeners. This score is identical to the percentage of correct responses for Dutch initial voiced stops produced without prevoicing found by Van Alphen and Smits (2004). As expected, the English listeners categorized all tokens of initial voiced stops accurately, regardless of the presence or absence of prevoicing.

Finally, for the /d/-/t/ contrast, no interaction was found between condition and native language (F (2, 76) = 1.18, $p > .1$). There was a significant effect of condition (F (2, 76) = 10.54, $p < .001$). A posthoc Bonferroni test showed that the sensitivity scores were higher in initial position (note that all initial /d/s were prevoiced) than in final position with release burst ($p < .05$) and in final position without release burst ($p < .001$), and that the two conditions in final position did not differ significantly from one another ($p > .1$). No effect of native language was found (F (1, 38) < 1). There was no effect of native language on bias for initial position (F (1, 39) = 1.08, $p > .1$), for final position with release burst (F (1, 39) = 1.63, $p > .1$), or for final position without release burst (F (1, 39) = 3.75, $p = .060$).

For both Dutch and English listeners, the removal of the release burst affected the categorization of the /b/-/p/ contrast but not of the /d/-/t/ contrast. The signal remaining after removal of the release burst may have contained clearer cues for the alveolar stops than for the bilabial stops. For example, the difference in vowel duration and in F1 offset frequency of /d/ and /t/ was larger than that of /b/ and /p/ (Table 4).

In general, the results for the four consonant contrasts show a consistent pattern. For the sensitivity measure, no interactions between position and native language were found for the contrasts /v/-/f/, /z/-/s/, and /d/-/t/. The Dutch listeners' categorization of /v/-/f/ and /z/-/s/ was as accurate in final position as in initial position, and as accurate as that of the English listeners. For the /d/-/t/ contrast, both groups performed better on initial position than on final position with release burst. For the /b/-/p/ contrast, there was an interaction between condition and native language. The English listeners outperformed the Dutch listeners on the initial position. The absence of prevoicing of initial stops hindered the Dutch listeners' categorization in English as much as it does in Dutch. As voiced stops are prevoiced less often in English than in Dutch, this may cause Dutch listeners to misperceive the voicing of initial stops in English more frequently than in Dutch. Dutch and English listeners performed equally well on the /b/-/p/ contrast in final position with release burst. Although Dutch does not allow for voiced obstruents in final, prepausal position, the Dutch listeners did not have a bias towards voiceless responses in final position.

The results are in line with the predictions of the PAM. Whereas the Dutch listeners categorized the /æ/-/ɛ/ contrast less accurately than the English listeners did, they categorized the English final voicing contrasts as accurately as the initial contrasts, and as accurately as the English listeners did.

Experiment 1 also tested categorization accuracy for final stops without a release burst. A difference between the Dutch and English listeners' categorization accuracy could have indicated a differential use of the duration of the preceding vowel as a voicing cue. However, the removal of the release bursts of final stops did not influence the Dutch and the English listeners differentially. For the /b/-/p/ contrast, Dutch and English listeners performed better on items with release burst than on the same tokens without release burst. The removal of the release burst affected the performance of the two language groups to the same extent. For the /d/-/t/ contrast there was no difference in sensitivity to items with or without release burst, neither for the Dutch nor for the English listeners. Thus, Experiment 1 did not provide any evidence that the Dutch listeners used vowel duration less than the English listeners did.

However, Experiment 1 was not a direct test of the use of vowel duration as a cue. Apart from vowel duration, several other cues remained available after removal of the release burst (e.g., F1 offset frequency, closure voicing; see Table 4). Thus, the Dutch listeners may have achieved a native-like level of accuracy without using vowel duration as a cue. On the other hand, the absence of a release burst may have stimulated the Dutch listeners to use vowel duration, while they may not do so when more perceptual cues are available. The results from this experiment are not decisive about these possibilities. In fact, any evidence of Dutch listeners using vowel duration as a cue for final voicing would leave open the possibility that their use of the cue was a reaction to the task at hand.

Therefore, the use of vowel duration as a cue to final obstruent voicing was investigated from a different angle in Experiment 2. In this experiment, stimulus materials were constructed such that they did not stimulate but rather discouraged the use of vowel duration as a voicing cue. The question was addressed whether Dutch listeners use vowel duration as a voicing cue as persistently as English listeners do. If the Dutch listeners did not use vowel duration in Experiment 2, this would not imply that they never do so. However, it could show that Dutch listeners do not use this cue as persistently as English listeners.

It was argued that among the languages that do not have voiced and voiceless obstruents in final position, Dutch prepares its listeners well for the use of vowel duration as a cue to English word-final obstruent voicing. As Dutch has long and short vowels, Dutch listeners are familiar with the use of phonetic vowel duration. They also have native-language experience with the use of vowel duration as a cue to word-medial obstruent voicing (Van den Berg, 1989). Although the role of vowel duration as a cue to voicing seems to be smaller in Dutch than in English, Dutch listeners' familiarity with the cue in word-medial position may facilitate its use in word-final position in English. Especially advanced learners of English may have learned to use this word-final voicing cue, combining their native and nonnative language experience. Therefore, Dutch listeners with a high level of proficiency in English might be expected to process the English obstruent voicing contrast in a native-like manner, with a native-like use of the vowel duration cue. If Dutch listeners do not use vowel duration in a native-like manner, however, this would raise the question whether nonnative listeners can ever be expected to process nonnative but familiar phoneme contrasts in unfamiliar positions in a native-like manner.

In Experiment 2, Dutch and English listeners' categorization of English final voiced and voiceless obstruents was investigated again. For reasons of comparison, categorization of initial voicing contrasts was tested as well. For practical reasons regarding the construction of phoneme continua, only fricatives were tested.

Experiment 2

Method

Participants

Twenty-eight native speakers of Dutch and 28 native speakers of British English, none of whom had participated in Experiment 1, took part in the experiment. The Dutch participants had a high level of proficiency in English as a second language (as in Experiment 1), whereas the English participants did not know any Dutch. The Dutch participants were recruited from the Max Planck Institute participant pool, and the English participants were recruited from the participant pool of the Laboratory of Experimental Psychology of the University of Sussex or at the University of Birmingham. None reported any hearing loss. All were volunteers and received a small fee for participation.

Materials

The same nonwords for initial and final fricative contrasts were used as in Experiment 1. The materials were recorded by the same native speaker of British English who recorded the materials for Experiment 1. The speaker read the items one by one, separated by a pause, in a clear citation style. The materials were recorded with a Sennheiser microphone in a soundproof booth onto digital audiotape and downsampled to 16 kHz during transfer to a computer. For each contrast, two target sounds and one or two carriers were extracted from the sound file, using the speech editor Xwaves. The target sounds were used to create voicing continua which were spliced onto the appropriate carriers, as described below.

From the nonword /zi:f/, /i:f/ was extracted, removing the initial /z/, with the cut being made at the first positive zero crossing after the offset of frication noise. From /fu:s/ and /fu:z/, /fu:/ was extracted, truncating the signal at the last positive zero crossing before the onset of frication noise. In a similar way, /u:k/ was extracted from /vu:k/, removing the initial /v/, and /ku:/ was extracted from /ku:f/ and /ku:v/, removing the final /f/ and /v/. These elements served as carriers.

An initial /z/ was extracted from another token of /zi:f/, truncating the signal at the first positive zero crossing after the offset of the frication noise. A final /z/ was extracted from another token of /fu:z/, with the cut being made at the last positive zero crossing before the start of the frication noise. Similarly, an initial and a final /s/, /v/, and /f/ were extracted from other tokens of /si:f/, /fu:s/, /vu:k/, /ku:v/, /fu:k/, and /ku:f/, respectively. For the initial and final /s/ and /f/, a portion in the center of the fricative was removed, such that the duration of each voiceless fricative matched the duration of its voiced counterpart. The initial /s/ and /f/ were shortened by 2 ms and 8 ms, respectively, and the resulting durations were 115 ms for the /s/ and 148 ms for the /f/. The final /s/ was shortened by 80 ms to 187 ms, and the final /f/ was shortened by 56 ms to 127 ms. The final /s/ and /f/ were shortened by 30 % and 31 %, respectively. The four pairs of phonemes that were thus obtained served as the endpoints of the four continua. For each continuum, nine intermediate steps were generated, following the procedure of Stevenson (1979) and Repp (1981). In this procedure, the amplitudes of two waveforms are added in varying proportions. The proportions have a ratio of 0-1 and 1-0 in the two endpoints, and are equally spaced in the intermediate steps, always adding up to 1.

For each continuum, the two endpoints and the nine intermediate steps were spliced onto the appropriate carriers. Thus, the resulting stimuli ranged from /zi:f/ to /si:f/ and from /fu:z/ to /fu:s/ for the alveolar fricatives, and from /vu:k/ to /fu:k/ and from /ku:v/ to /ku:f/ for the labiodental fricatives. For the two initial continua, there was one carrier each. For the two final continua, there were two carriers each. One was originally pronounced with a voiceless final fricative and contained a phonetically short vowel (of 118 ms for the /z/-/s/ contrast and 98 ms for the /v/-/f/ contrast), the other was originally pronounced with a voiced final fricative and contained a phonetically long vowel (of 233 ms for the /z/-/s/ contrast and 257 ms for the /v/-/f/ contrast).

Design

Stimuli were blocked by contrast, position, and carrier. Each block was presented to half of the participants. As there were two carriers for final contrasts and one for initial contrasts, half of the participants only heard the two final contrasts, and the other half heard the two final contrasts and the two initial contrasts. The order of the blocks with initial and final contrasts was counterbalanced where applicable. As explained above, the non-target consonants in the CVC items were obstruents. As this restriction yields a limited number of items that are nonwords in both languages, /f/ was part of the carrier in the items where /z/

and /s/ were the target sounds. Therefore, the /v/-/f/ contrast was always tested before the /z/-/s/ contrast. Each block consisted of 20 repetitions of the 11 steps of the continuum, semi-randomized such that the same step could not occur twice in succession.

Crucially, each participant was presented with only one carrier for each final contrast. For each participant, the duration of the vowel for each final contrast was unvarying. Thus, vowel duration was not informative for the voicing contrast. For all participants, there was a mismatch between vowel duration and other information in the signal for a subset of the stimuli. When voiced fricatives were preceded by a short vowel, or voiceless fricatives by a long vowel, vowel duration and information in the frication noise pointed in opposite directions.

Procedure

The procedure was as described for Experiment 1. Each block was preceded by a practice part containing two presentations of each of the 11 steps of the continuum in semi-randomized order. The response buttons were labeled “Z” and “S”, or “V” and “F”, respectively.

Results and discussion

Eight responses with RTs longer than 10,000 ms due to a technical error were removed from the analysis. The categorization curves of each contrast in each position and for each subject separately were fitted with logistic regression. From the regression models, 50-percent crossover points were retrieved, reflecting the location of the category boundary. From the models, a measure of the steepness of the categorization curve at the 50-percent crossover point was computed, indicating how categorical perception was. In five response sets, the percentage of correct responses at step 1 or 11 did not exceed 50 % (one response set representing one contrast in one position for one subject). No logistic regressions were performed on those response sets.

The categorization results for the initial contrasts are presented in Table 7. There were no differences between the Dutch and English listeners in the steepness of the slopes, either for the /z/-/s/ contrast (F (1, 27) = 1.55, $p > .1$) or for the /v/-/f/ contrast (F (1, 26) < 1).

The categorization results for the final /z/-/s/ contrast are presented in Figure 1. If vowel duration was used for the categorization of ambiguous fricatives from the middle region of the voicing continuum, this should have resulted in a shift between the curves corresponding to the long and short vowel conditions. However, an ANOVA on the 50-percent crossover points showed no interaction between vowel duration and native language (F (1, 54) = 2.53, $p > .1$), and no main effects of vowel duration (F (1, 54) = 1.17, $p > .1$) or native language (F (1, 54) < 1).

Table 7. Experiment 2 results: Mean percentage of "z" or "v" responses to initial fricatives as a function of the place on an 11-step stimulus continuum ranging from /z/ to /s/ or from /v/ to /f/ and participants' native language.

	Dutch		**English**	
	/z/-/s/	**/v/-/f/**	**/z/-/s/**	**/v/-/f/**
1 (voiced)	85	93	95	98
2	83	94	91	96
3	83	94	94	97
4	83	93	92	93
5	83	87	91	88
6	79	72	83	76
7	63	53	78	55
8	56	37	65	27
9	41	23	50	12
10	28	15	26	6
11 (voiceless)	10	9	8	3

The categorization results for the final /v/-/f/ contrast are presented in Figure 2. The graphs show that vowel duration had a differential effect on the Dutch and the English participants' categorization results. A significant interaction was found between the effects of vowel duration and native language on 50-percent crossover point (F (1, 52) = 4.32, p < .05).

For the Dutch listeners, the categorization curves were similar in the conditions with the short and with the long preceding vowel. Although the curve for items with a long vowel was located slightly further towards the voiceless side of the continuum than the curve for the short vowel, the 50-percent crossover points were not statistically different (F (1, 27) < 1).

For the English listeners, the categorization curve for the items with a long vowel was strongly shifted relative to the curve for items with a short vowel. The 50-percent crossover points were significantly different in the two conditions (F (1, 24) = 18.63, p < .001), with a larger 50-percent crossover point for the condition with longer vowel duration, showing a preference for "v" responses which persisted further towards the voiceless side of the continuum.

Moreover, the English listeners categorized even the /f/ endpoint as "v" 31 % of the time when it was preceded by a long vowel. The difference between the curves for the short and long vowel conditions was located on the voiceless side of the continuum. ANOVAs on arcsine-transformed proportions showed that from steps 6 to 11 the proportion of "v" responses was significantly higher for items with a long vowel than for items with a short vowel (p < .01).

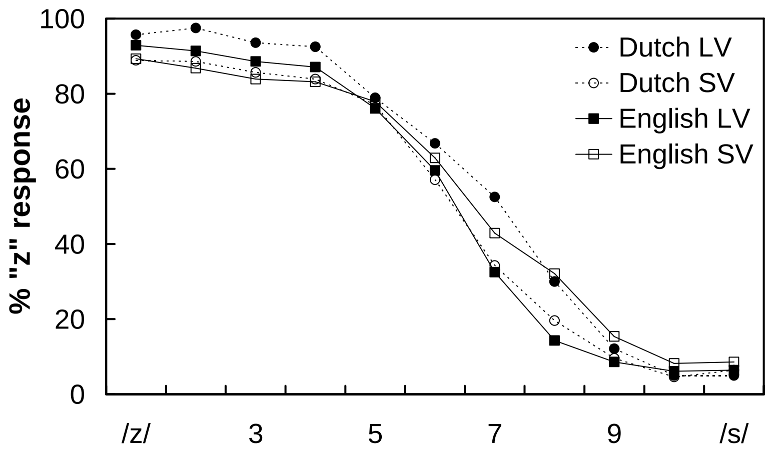

Figure 1. Experiment 2: Mean percentage of "z" responses to final fricatives as a function of the place on an 11-step stimulus continuum ranging from /z/ to /s/, preceding vowel duration (LV: long vowel; SV: short vowel), and participants' native language.

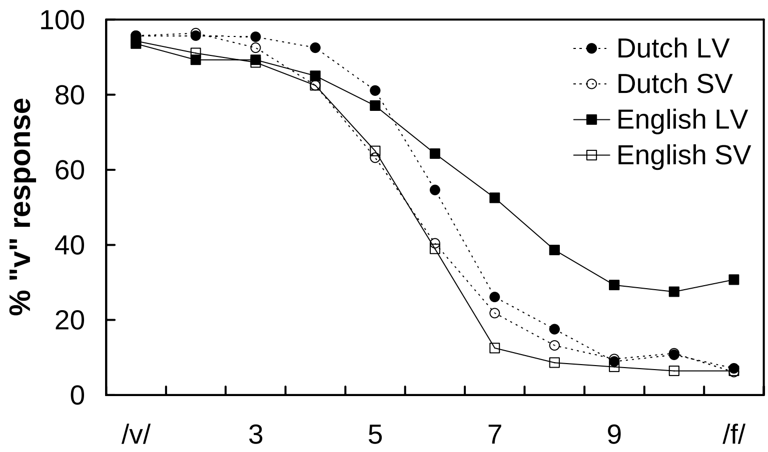

Figure 2. Experiment 2: Mean percentage of "v" responses to final fricatives as a function of the place on an 11-step stimulus continuum ranging from /v/ to /f/, preceding vowel duration (LV: long vowel; SV: short vowel), and participants' native language.

There was no effect of native language on steepness of the slope in the short vowel condition (F (1, 27) < 1). In the long vowel condition, the Dutch listeners' categorization curve was steeper than the English listeners' curve (F (1, 24) = 4.76, p < .05), indicating that the Dutch listeners' categorization was more categorical than the native English listeners' categorization.

Vowel duration thus affected the categorization of the /v/-/f/ contrast and the /z/-/s/ contrast differentially. For the final /v/-/f/ contrast, a change in vowel duration led to a shift in the categorization curve for the English but not for the Dutch listeners. For the final /z/-/s/ contrast, there was no shift for either language group. A significant three-way interaction (F

(1, 48) = 10.12, $p < .001$) among the effects of vowel duration, native language, and place of articulation on 50-percent crossover point confirmed that the contrasts differed in this respect.

The finding that there was no effect of vowel duration for the final /z/-/s/ contrasts is not surprising in itself, as vowel duration was not informative in this experiment. As vowel duration was kept constant for each participant throughout the whole block, it did not have any cue value for the voicing contrast. Nevertheless, the English listeners but not the Dutch listeners showed an effect of vowel duration in their categorization of the final /v/-/f/ contrast. As the English listeners based their categorization decisions on the uninformative vowel duration, their categorization for the /v/-/f/ contrast preceded by a long vowel was less categorical than the Dutch listeners' categorization.

The different results for the alveolar and labiodental fricatives may be caused by their acoustic characteristics. Word-initial alveolar fricatives have a higher amplitude than labiodental fricatives (Jongman, Wayland, & Wong, 2000). A similar difference may exist in final position. Indeed, the spectral power of the final /z/ and /s/ was higher than that of the final /v/ and /f/ in Experiment 1 (Table 3) and in Experiment 2 (where the mean logarithms of the spectral power of the frication noise above 500 Hz in Pa^2 were -3.1 and -4.3, respectively). Further, in final position, alveolar fricatives have a longer noise duration than labiodental fricatives (Crystal & House, 1988). This was also the case in Experiment 1 (Table 3) and in the original final fricatives in Experiment 2. Therefore, the information in the alveolar frication signal may generally be more easily perceptible for the listener than the information in a labiodental. A less informative frication signal may stimulate listeners to exploit other sources of information. This may explain why the English participants took vowel duration into consideration in their decisions for the final /v/-/f/ contrast, but not for the /z/-/s/ contrast.

English listeners categorized the same tokens significantly more often as "v" when they were preceded by a phonetically long vowel than when they were preceded by a short vowel. Even the endpoint /f/ was categorized as "v" 31 % of the time in the long vowel condition. Apparently, the long duration of the vowel pointed towards a voiced fricative so strongly that it overruled the other information in the signal in many cases. Note that the reverse did not happen on the other side of the continuum: tokens at the voiced end of the continuum received a high percentage of voiced responses, even when preceded by a short vowel. Several factors may have contributed to this asymmetry. In the first place, the final /f/ was shortened to match the duration of the final /v/. As frication duration is a cue to voicing (e.g., Watson, 1983), the shortening made the final /f/ more /v/-like. Note that this shortening did not lead to a high percentage of "v" responses for the Dutch listeners, or for the English listeners in the short vowel condition. Neither was there a high percentage of "z" responses to the endpoint /s/, which was shortened to a similar extent. Thus, the shortening cannot explain the high percentage of "v" responses for the English listeners in the long vowel condition, but it may have made the /f/ endpoint more acceptable as a “v” than vice versa. Second, listeners have experience with vowel shortening in fast speech. In fast speech, vowels are reduced relatively more than consonants (Gay, 1978), so that not only the vowel duration itself, but also the ratio of vowel and fricative duration changes. Indeed, vowel lengthening before voiced obstruents decreases at faster speaking rates (Smith, 2002). As a result of this experience with absolute and relative vowel shortening, listeners may find it easier to ignore short vowel duration as a cue to voicing when it mismatches with other cues than to ignore long vowel duration. In the third place, phonologically voiced obstruents are phonetically often unvoiced (Stevens, Blumstein, Glicksman, Burton, & Kurowski, 1992). Therefore, listeners may show asymmetric weighting of the presence or absence of phonetic voicing. The presence of phonetic voicing may signal a voiced obstruent relatively strongly, while its absence may not point as strongly towards a voiceless interpretation. This may have contributed to the finding that tokens on the voiced side of the continuum were predominantly perceived as voiced, whereas in the long vowel condition, tokens on the voiceless side of the continuum were less often perceived as voiceless.

General discussion

The results of the two experiments presented in this paper show that a native-like level of accuracy may be reached for the categorization of nonnative phonemes, even though the phonemes are not necessarily processed in a native-like manner.

Experiment 1 investigated the accuracy with which Dutch listeners categorized English contrasts with different degrees of correspondence in Dutch phonology. The /æ/-/ɛ/ contrast, which the PAM predicts to be among the most difficult English contrasts for Dutch listeners, was indeed found to be the most difficult contrast in this experiment. Although the Dutch listeners performed amply above chance, the English listeners showed a significantly higher sensitivity than the Dutch listeners did. The obstruent voicing contrasts are matched by similar contrasts in Dutch, and the PAM predicts them to be easy to distinguish. Although Dutch voicing contrasts do not occur in final position, the Dutch listeners categorized the English final voicing contrasts as accurately as (or even more accurately than) the initial contrasts, and as accurately as the English listeners did. Dutch listeners were not biased towards voiceless responses in final position.

The PAM does not make any predictions about the perception of familiar contrasts in unfamiliar positions. The present study suggests that an unfamiliar position does not necessarily complicate the perception of familiar but nonnative contrasts. For example, in Experiment 1 the Dutch listeners categorized the /b/-/p/ contrast more accurately in the unfamiliar final position than in the familiar initial position. In order to make predictions about the perception of familiar contrasts in familiar and unfamiliar positions, it is important to take into account that speech sounds have different acoustic characteristics in different positions. The extent to which these characteristics overlap with those of the native speech sounds seems an important predictor of the ease with which nonnative listeners can distinguish between the sounds.

In Experiment 1, no evidence was found that the Dutch listeners used vowel duration as a cue to final voicing less than the English listeners did. The removal of the release burst from final stops did not affect the Dutch listeners' categorization more than it affected the English listeners' categorization. However, the Dutch listeners may have achieved a native-like level of accuracy without using vowel duration as a cue. Flege (1989) found a native-like level of accuracy for Chinese listeners' categorization of unedited tokens of English final /d/-/t/. Nonetheless, the Chinese listeners were found to rely on cues in the release burst more than the English listeners did. They had achieved a native-like accuracy through a nonnative-like manner of processing. When the release burst was removed, they were no longer able to maintain a native-like level of accuracy. The Dutch listeners may have been better able to adapt to the removal of the release burst. Their knowledge about English voicing cues may have been sufficient to use those cues that remained available when the burst had been removed (e.g., F1 offset frequency, closure voicing). They may have achieved a native-like level of accuracy for the categorization of stops without release burst without using vowel duration. On the other hand, it is also possible that the absence of a release burst stimulated the Dutch listeners to use vowel duration, while they may not do so when more perceptual cues are available. Thus, the results from Experiment 1 left the possibility open that nonnative listeners do not need to process nonnative phonemes in a native-like manner in order to achieve a native-like level of accuracy.

Therefore, Experiment 2 tested the use of the duration of the preceding vowel as a cue to final fricative voicing with items which were constructed such that they did not stimulate the use of vowel duration as a voicing cue. In Experiment 2, categorization of initial fricative voicing contrasts was tested as well. In line with the results from Experiment 1, no differences were found between Dutch and English listeners' categorization of initial fricative voicing contrasts.

In order not to stimulate the Dutch listeners to use vowel duration more than they would normally do, but rather discourage its use, the vowel duration cue was kept uninformative. Vowel duration even mismatched with other information in the signal for some of the tokens. Nevertheless, the English listeners tried to use vowel duration for the categorization of /v/ and /f/. Especially, the categorization of tokens with long vowels was often consistent with vowel duration. Apparently, for the English listeners, vowel duration was such an important cue for final voicing that it often overruled other information in the signal. The Dutch listeners, on the other hand, did not use vowel duration at all. Even in the middle range of the continuum, for tokens with an ambiguous identity, there was no effect of vowel duration on the Dutch listeners' categorization of final /z/-/s/ or /v/-/f/. Thus, in Experiment 2, Dutch listeners did not categorize final voiced and voiceless obstruents in a native-like manner. The Dutch listeners were able to ignore vowel duration when it was uninformative and misleading. They differed in this respect from the English listeners, who ignored vowel duration for the categorization of the final /z/-/s/ contrast, but relied heavily on it for the categorization of the final /v/-/f/ contrast. As a result, for the /v/-/f/ contrast preceded by a long vowel, Dutch listeners' categorization curve was steeper than that of the native English listeners. As vowel duration was not informative and sometimes mismatched with other voicing cues, English listeners' use of vowel duration for the /v/-/f/ contrast resulted in less categorical perception.

The Dutch listeners, who did not use vowel duration as a cue in Experiment 2, may do so in other circumstances where the cue is informative. Indeed, they may have used vowel duration as a cue in Experiment 1. However, the results from Experiment 2 showed that the Dutch listeners did not use vowel duration as persistently as the English listeners did. This may be a result from their native-language experience, where vowel duration is a less important cue to (word-medial) obstruent voicing than in English. From their native-language experience, the Dutch listeners may have inferred that vowel duration is only a minor cue to English final obstruent voicing as well. Another explanation could be that Dutch listeners are regularly exposed to English spoken by native speakers of Dutch. Elsendoorn (1985) has shown that the difference in vowel duration before voiced and voiceless final obstruents in the English spoken by Dutch learners is smaller than that in the English of native speakers. From exposure to English spoken by Dutch learners, Dutch listeners may have learned to ignore vowel duration as a voicing cue when it is uninformative.

It was argued that Dutch prepares its listeners well for the distinction of English word-final obstruent voicing. Dutch has obstruent voicing contrasts which are perceptually similar to English contrasts, which makes the English contrasts easy to distinguish according to the PAM (Best et al., 1988). Dutch allows for voiceless obstruents in word-final position, which has been found to facilitate perception of the distinction (Flege & Wang, 1989). Dutch provides experience with phonetic vowel duration for the distinction of phonemically long and short vowels, which has been found to facilitate the use of this cue for the final consonant voicing contrast (Crowther & Mann, 1992), and Dutch provides experience with the use of vowel duration as a cue to word-medial obstruent voicing (Van den Berg, 1989). Indeed, Dutch listeners were found to categorize English final obstruent voicing with a native-like level of accuracy. Nevertheless, they were found to use vowel duration as a cue less persistently than the English listeners did.

As even native listeners of Dutch, who had the necessary experience for a native-like use of vowel duration, and who had a high level of proficiency in English, did not use perceptual cues in a native-like manner, this raises the question whether nonnative listeners can ever be expected to process nonnative but familiar phoneme contrasts in unfamiliar positions in a native-like manner.

Although the Dutch listeners in this study had a high level of proficiency in English, they had not reached a level of ultimate attainment. Their English perception skills were still open to improvement. However, such improvement may not involve the use of perceptual cues for contrasts which they could already accurately distinguish. As the results from the experiments in this paper show, a native-like level of accuracy can be achieved even when the perceptual cues are not processed in a native-like manner. Possibly, Dutch listeners can learn to use vowel duration as a cue to English word-final obstruent voicing in a fully native-like manner with a native-like persistence, for example through laboratory-based training, but there may be no need to learn this for normal language use. Presumably, the second language learner's goal is not to process language in a native-like manner, but rather to be able to understand (and produce) the language well enough to meet the learner's communicative needs. If a native-like use of perceptual cues is not necessary for accurate perception, it is possible that listeners may never learn to perceive nonnative phonemes in a native-like manner. The benefits of native-like processing may be too small, or even nonexistent.

It should be relatively easy for Dutch listeners to learn to use vowel duration in an English native-like manner, and nonetheless, in Experiment 2 the Dutch listeners did not use vowel duration as a cue when the English listeners did. As even these Dutch listeners did not use vowel duration in a native-like way, it seems likely that native listeners of languages for whom the cue is harder to learn would in many cases not use the cue in a native-like manner either. The reasoning can be extended to other perceptual cues as well. The vowel duration cue provides a great amount of information about the English voicing contrast, and it is a very important cue for English listeners. It therefore seems a good candidate for native-like use by nonnative listeners. As even this cue was not used in a native-like way, it seems likely that other, less informative cues would in many cases not be used in a native-like manner either. Thus, for the distinction of nonnative but familiar contrasts in unfamiliar positions, a native-like manner of phonetic processing may hardly ever be attained.

References

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

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.

Best, C. T., McRoberts, G. W., & Sithole, N. M. (1988). Examination of perceptual reorganization for nonnative speech contrasts: Zulu click discrimination by English-speaking adults and infants. *Journal of Experimental Psychology: Human Perception and Performance, 14*, 345-360.

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

Booij, G. (1995). *The Phonology of Dutch*. Oxford: Oxford University Press.

Byrd, D. (1993). 54,000 American stops. *UCLA Working Papers, 83*, 97-115.

Crowther, C. S., & Mann, V. (1992). Native language factors affecting use of vocalic cues to final consonant voicing in English. *Journal of the Acoustical Society of America, 92*, 711-722.

Crystal, T. H., & House, A. S. (1988). A note on the durations of fricatives in American English. *Journal of the Acoustical Society of America, 84*, 1932-1935.

Elsendoorn, B. A. G. (1985). Production and perception of Dutch foreign vowel duration in English monosyllabic words. *Language and Speech, 28*, 132-254.

Flege, J. E. (1989). Chinese subjects' perception of the word-final English /t/-/d/ contrast: Performance before and after training. *Journal of the Acoustical Society of America, 86*, 1684-1697.

Flege, J. E., & Hillenbrand, J. (1987). A differential effect of release bursts on stop voicing judgments of native French and English listeners. *Journal of Phonetics, 15*, 203-208.

Flege, J. E., & Wang, C. (1989). Native-language phonotactic constraints affect how well Chinese subjects perceive the word-final English /t/-/d/ contrast. *Journal of Phonetics, 17*, 299-315.

Gay, T. (1978). Effect of speaking rate on vowel formant movements. *Journal of the Acoustical Society of America, 63*, 223-230.

Jongman, A., Sereno, J. A., Raaijmakers, M., & Lahiri, A. (1992). The phonological representation of [voice] in speech perception. *Language and Speech, 35*, 137-152.

Jongman, A., Wayland, R., & Wong, S. (2000). Acoustic characteristics of English fricatives. *Journal of the Acoustical Society of America, 108*, 1252-1263.

Kuijpers, C. T. L. (1996). Perception of the voicing contrast by Dutch children and adults. *Journal of Phonetics, 24*, 367-382.

Lisker, L., & Abramson, A. S. (1964). A cross-language study of voicing in initial stops: Acoustical measurements. *Word, 20*, 384-422.

MacMillan, N. A., & Creelman, C. D. (1991). *Detection Theory: A User's Guide*. Cambridge: Cambridge University Press.

McNicol, D. (1972). *A Primer of Signal Detection Theory*. Sydney: Australasian Publishing Company.

Peterson, G. E., & Lehiste, I. (1960). Duration of syllable nuclei in English. *Journal of the Acoustical Society of America, 32*, 693-703.

Raphael, L. J. (1972). Preceding vowel duration as a cue to the perception of the voicing characteristic of word-final consonants in American English. *Journal of the Acoustical Society of America, 51*, 1296-1303.

Repp, B. H. (1981). Perceptual equivalence of two kinds of ambiguous speech stimuli. *Bulletin of the Psychonomic Society, 18*, 12-14.

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.

Slis, I. H., & Cohen, A. (1969a). On the complex regulating the voiced-voiceless distinction I. *Language and Speech, 12*, 80-102.

Slis, I. H., & Cohen, A. (1969b). On the complex regulating the voiced-voiceless distinction II. *Language and Speech, 12*, 137-155.

Slis, I. H., & Van Heugten, M. (1989). Voiced-voiceless distinction in Dutch fricatives. In H. Bennis & A. van Kemenade (Eds.), *Linguistics in The Netherlands 1989* (pp. 123-132). Dordrecht, The Netherlands: Foris.

Smith, B. L. (1978). Effects of place of articulation and vowel environment on 'voiced' stop consonant production. *Glossa, 12*, 163-175.

Smith, B. L. (2002). Effects of speaking rate on temporal patterns of English. *Phonetica, 59*, 232-244.

Stevens, K. N., Blumstein, S. E., Glicksman, L., Burton, M., & Kurowski, K. (1992). Acoustic and perceptual characteristics of voicing in fricatives and fricative clusters. *Journal of the Acoustical Society of America, 91*, 2979-3000.

Stevenson, D. C. (1979). *Categorical Perception and Selective Adaptation Phenomena in Speech.* Unpublished doctoral dissertation, University of Alberta, Edmonton, Canada.

Trubetzkoy, N. S. (1977). *Grundzuege der Phonologie* (original work published 1939) (6th ed.). Goettingen, Germany: Van den Hoeck & Ruprecht.

Van Alphen, P. M., & Smits, R. (2004). Acoustical and perceptual analysis of the voicing distinction in Dutch initial plosives: The role of prevoicing. *Journal of Phonetics, 32*, 455-491.

Van den Berg, R. J. H. (1989). Perception of voicing in Dutch two-obstruent sequences: Covariation of voicing cues. *Speech Communication, 8*, 17-25.

Watson, I. (1983). Cues to the voicing contrast: A survey. Cambridge Papers in Phonetics and Experimental Linguistics, 2.

Spurious lexical activation in nonnative listening

Chapter 3

Broersma, M. (submitted). Spurious lexical activation in nonnative listening. Cognition.

Abstract

Spurious lexical activation in nonnative listening was investigated with lexical decision, cross-modal priming, and phonetic categorization experiments. 'Near-words' were used, which differed from real words in one phoneme contrast (e.g., *deaf* – *daf*, *globe* - *glope*). *Deaf* from *DEFinite* activated the lexical representation of *deaf* for Dutch and English listeners alike, whereas *daf* from *DAFfodil* activated *deaf* for Dutch listeners but not for English listeners. Similarly, *lemp* from *eviL EMPire* activated *lamp* for Dutch listeners only. For nonnative listeners, near-words often caused as much activation as words did. The stimuli *deaf* and *daf* were perceptually ambiguous for Dutch listeners, *globe* and *glope* were not. Both perceptually ambiguous and unambiguous near-words caused more lexical activation for the Dutch listeners than for the English listeners. Thus, lexical competitors which are not active in native listening are active in nonnative listening, even when phoneme perception is uncompromised.

Introduction

When people listen to speech, they are confronted with many unintended words. Although listeners may not be aware of it, a speech signal may temporarily match many different word forms. A large body of evidence has established that words in the mental lexicon are activated when they match the acoustic input, even if the match is only partial. Zwitserlood (1989), for example, demonstrated that multiple words which are partially consistent with the input are simultaneously activated. When listeners heard the first part of either the Dutch word *kapitein* (‘captain’) or *kapitaal* (‘capital’), both word forms were activated. Thus, a temporary partial match with the input leads to lexical activation. But the speech signal may also more fully match word forms which are nevertheless unintended by the speaker. Speech may contain words which are embedded in the signal, either within one word or spanning a word boundary. There is ample evidence that these embedded words activate their corresponding representation in the mental lexicon.

For example, when listeners hear the word *captain*, not only the word *captain* but also the word *cap* is temporarily activated, and hearing *cap* leads to the activation of both *cap* and *captain* (Davis, Marslen-Wilson, & Gaskell, 2002). Similarly, Salverda, Dahan, and McQueen (2003) showed that presentation of *hamster* caused activation of both the words *hamster* and *ham*. In a series of eye-tracking experiments they further showed that durational information was used to favor the correct interpretation of an ambiguous sequence. A longer duration of the first syllable favored the activation of the embedded word *ham*, and a shorter duration favored the carrier word *hamster*. However, this bias was small and did not prevent activation of the other word.

Lexical representations may not only be activated when they match with the beginning of a longer word, but also when they correspond to the last part of a word. When listeners hear *trombone*, the word *bone* is also activated. Shillcock (1990) found no difference in the level of activation of *bone* after presentation of *trombone* or *bone*. Vroomen and De Gelder (1997) found similar results for Dutch. They showed that embedded words were activated when they corresponded to the final syllable of a disyllabic carrier word, either with a strong - strong or a weak - strong pattern. Thus, *framboos* (‘raspberry’) activated *boos* (‘angry’), and *beschuit* (‘biscuit’) activated *schuit* (‘boat’). No activation was found for words which were embedded at the end of a monosyllabic word, like *wijn* (‘wine’) in *zwijn* (‘swine’).

Unintended words may also be found across word boundaries. Gow and Gordon (1995) showed that the ambiguous sequence *two lips* activated both the word *lips* and *tulips*. The sequence *two lips* activated the word form *tulips* as strongly as the input *tulips* did. Tabossi, Burani, and Scott (1995) found similar results for Italian. They showed that a sentence containing the sequence *visi tediati* ('faces bored') at least temporarily activated the trisyllabic word *visite* ('visits') as strongly as a sentence containing the word *visite* did.

The activation of unintended words may come to an end when a mismatch occurs between the input and the lexical representation. For example, whereas *kapit-* activated both *kapitein* and *kapitaal*, the activation of the incorrect interpretation decreased when disambiguating information was presented, and by the time the whole word had been heard, no activation of the incorrect interpretation remained (Zwitserlood, 1989).

Thus, there is extensive evidence that multiple lexical candidates are activated during speech comprehension. Multiple lexical activation is a central assumption that all current models of speech comprehension share (for a review see McQueen, 2004). In all models, the amount of lexical activation depends on the extent to which the speech signal and the lexical representation match. A mismatch between the input and the lexical representation leads to the (partial) deactivation of the lexical representation. However, the models differ in the way in which this deactivation is achieved.

The Cohort model and Shortlist propose that mismatches lead to bottom-up inhibition of lexical representations. In the original Cohort model (Marslen-Wilson & Welsh, 1978), a mismatch leads to the removal of a lexical representation from the competitor set. In a more recent version of the Cohort model (Marslen-Wilson, 1987) a mismatching lexical representation is not excluded from the competition process, but its activation level decreases. Similarly, in Shortlist (Norris, 1994) a mismatching lexical representation is penalized through bottom-up inhibition. Frauenfelder, Scholten, and Content (2001) present evidence for the occurrence of bottom-up inhibition.

In TRACE (McClelland & Elman, 1986) and Shortlist deactivation of mismatching lexical representations takes place through lateral inhibition at the lexical level. Lexical representations which are simultaneously active compete for recognition, and the most active representation deactivates the weaker ones. In Shortlist, bottom-up and lateral inhibition are thus combined to account for the recognition of the correct word and the suppression of the activation of partially matching competitors. A phoneme in the input which mismatches with the competitor decreases the activation of the competitor, while it adds to the activation of the fully matching target. Thereupon, the target word form is more active than the competitor and further deactivates the competitor through lateral inhibition.

Lateral inhibition may lead to the deactivation of competitors and to the selection of the intended word form. Thus, it contributes to the correct interpretation of the speech signal. However, it may also hinder recognition. As all activated word forms participate in the competition process, lateral inhibition may diminish the activation of the intended word. It is more difficult to recognize the word *sack* in the non-word stimulus *sacrif*, which is consistent with the competitor *sacrifice*, than in *sacrick*, which is not a word onset (McQueen, Norris, & Cutler, 1994).

The number of lexical competitors influences the ease with which a word is recognized. There is ample evidence that it is harder to recognize a word when more competitors are active. Vroomen and De Gelder (1995) tested the activation of Dutch words in a context consistent with different numbers of competitor words. For example, activation of the word *melk* ('milk') was investigated after presentation of the non-word *melkem*, *melkeum*, or *melkaam*. The competitor sets for the three types of stimuli had different sizes. There are no Dutch words beginning with *ke-* (/kə/), few with *keu-*, and many with *kaa-*. Activation of *melk* was found to be strongest after presentation of *melkem* and weakest after *melkaam*. The ease with which a word was recognized depended on the number of competitors beginning with the last phoneme of the word. Norris, McQueen, and Cutler (1995) found similar results for English. Furthermore, the number of lexical neighbors influences the ease with which a word can be recognized. The presentation of a word with a dense lexical neighborhood leads to the activation of a large amount of competitors, which has been found to hinder the recognition of the target word (for a review see Luce, Pisoni, & Goldinger, 1990; Vitevitch & Luce, 1999). Thus, the more lexical competitors become active, the harder it is to recognize the intended word.

As the activation of competitors renders the selection of the intended word more difficult, so the deactivation of competitors mismatching the input benefits speech comprehension. As a mismatch between the speech signal and a lexical representation leads to a decrease in activation of the lexical representation of the mismatching word, no unnecessary lexical competition is expected in the perception of the listener's native language. However, this may be different when people listen to a nonnative language. In nonnative listening, a segmental mismatch between the speech signal and a word may not always be perceived as a mismatch. Indeed, several studies have shown that there is more activation of lexical competitors in nonnative listening than in native listening.

First, there is evidence that minimal pairs sometimes activate each other more for nonnative than for native listeners. Pallier, Colomé, and Sebastián-Gallés (2001) tested highly fluent early Spanish-Catalan bilinguals whose native language was Spanish on their perception of Catalan minimal pairs which differed in contrasts that were difficult for these listeners to distinguish. In an auditory repetition-priming task, presentation of one word facilitated the recognition of the other word for these listeners, but not for listeners whose native language was Spanish. Broersma (2002, submitted) studied the perception of British English by Dutch listeners with a high level of proficiency in English as a second language. In a cross-modal priming task, listeners were presented with minimal pairs which differed in the vowels /æ/ and /ɛ/, which are difficult to distinguish for Dutch listeners (e.g., *flash – flesh*). Although the Dutch listeners did not always treat the words as homophones, presentation of one word led to the activation of the other word more often than for the native listeners. Cutler and Otake (2004) present evidence from auditory repetition priming that English minimal pairs differing in the /r/-/l/ contrast (e.g., *write – light*) activated one another for Japanese listeners, and minimal pairs differing in the /æ/-/ɛ/ contrast (e.g., *cattle – kettle*) for Dutch listeners.

Second, partially overlapping competitors may remain active longer in nonnative listening than in native listening. In a study of Dutch listeners' perception of English, Weber and Cutler (2004) used word pairs with onsets which overlapped except for contrasts that were difficult for these listeners to distinguish. In an eye-tracking experiment, after presentation of a word like *panda*, activation of its partially overlapping competitor *pencil* was found for the Dutch listeners but not for English listeners. Similarly, in a cross-modal priming experiment, Broersma (submitted) found that *daffo* from *daffodil* activated *deficit* more for Dutch listeners than for English listeners. In an eye-tracking experiment studying Japanese listeners' perception of English word pairs with onsets that overlapped except for the difficult to distinguish /r/-/l/ contrast, words like *rocket* were found to activate competitors like *locker* (Cutler, Weber, & Otake, in press).

Finally, there is some evidence that non-words which are embedded in the speech signal may cause more lexical activation for nonnative than for native listeners. Broersma (2002) first showed that non-words which differed from real English words in one phoneme caused more activation of the word form for Dutch listeners than for native English listeners. In a similar study, Sebastián-Gallés, Echeverría, and Bosch (2005) studied the perception of Catalan by highly fluent Spanish-Catalan early bilinguals whose native language was either Spanish or Catalan. Non-words were created by replacing the vowels /e/ and /ɛ/, which are difficult to distinguish for Spanish-dominant listeners, in Catalan words. In an auditory lexical decision experiment, the early Spanish-dominant bilinguals differentiated less between the non-words and the words than the Catalan-dominant bilinguals did.

If non-words cause more lexical activation for nonnative listeners than for native listeners indeed, this could severely complicate nonnative speech comprehension. Both native and nonnative listeners are regularly confronted with unintended words, which are plentiful in speech (e.g., McQueen, Cutler, Briscoe, & Norris, 1995). When *equaL AMPlitude* activates the word *lamp*, this word will compete for recognition with the target words *equal* and *amplitude*. Nonnative listeners may be confronted with the additional problem of 'near-words', which mismatch with lexical representations for native listeners, but may match with them for nonnative listeners. The sequence *eviL EMPire* will not provide a full match with *lamp* for native listeners of English, but it may do so for nonnative listeners. The presence of near-words in the signal may lead to a larger number of lexical competitors being active for nonnative listeners than for native listeners.

Near-words have a high frequency of occurrence in speech. Cutler (2005) computed the number of near-words that listeners may encounter in British English. Near-words differed from real words in the /æ/-/ɛ/ contrast or in the /r/-/l/ contrast. First, in real words, /æ/ was replaced with /ɛ/ and vice versa, and lexical statistics were computed to determine how often the resulting near-words occurred embedded in other words. The frequency of the near-words was considerable, with more than 78,000 occurrences per million words. Replacement of /r/ and /l/ yielded even more near-words, with more than 119,000 occurrences per million words.

An important question is what causes increased competitor activation in nonnative listening. A likely cause is perceptual ambiguity of nonnative speech sounds. There is overwhelming evidence that the perception of particular phoneme contrasts can be very difficult for nonnative listeners (e.g., Strange, 1995). Connine, Blasko, and Wang (1994) found that if the speech input corresponded equally well to different lexical representations, both representations remained active. For example, presentation of a phoneme which was ambiguous between /d/ and /t/ followed by *-ent* led to the activation of the words *dent* and *tent* to a comparable degree. Neither one of the interpretations outweighed the other representation enough to deactivate it and to end the competition. A similar situation may arise if nonnative phonemes are ambiguous for a listener. For example, if a listener does not hear the difference between the word *lamp* and the near-word *lemp*, either one will activate the lexical representation of *lamp*. Indeed, in the studies described above, minimal pairs (Broersma, 2002, submitted; Cutler & Otake, 2004; Pallier et al., 2001), partially overlapping words (Broersma, submitted; Cutler et al., in press; Weber & Cutler, 2004) and near-words (Broersma, 2002; Sebastián-Gallés et al., 2005) were based on phoneme contrasts which were perceptually ambiguous for the nonnative listeners.

However, Broersma (submitted) found that increased competitor activation also occurred for minimal pairs which differed in a contrast which was easy to distinguish for the nonnative listeners. Broersma (submitted) proposed that the nonnative listeners might have disregarded these word-final contrasts for lexical access, even though they were easy to distinguish, because the phoneme pairs were not contrastive in word-final position in their native language. As another possible explanation Broersma (submitted) proposed that lexical representations of nonnative listeners may sometimes differ from native listeners' representations. Especially if listeners regularly hear their second language spoken by other nonnative speakers, nonnative-like pronunciations may be stored in their mental lexicon. If a near-word matches such a pronunciation, it would effectively activate the lexical representation.

This paper further investigates the occurrence of spurious lexical activation for native listeners of Dutch during the comprehension of British English speech. Dutch listeners with a high level of proficiency in English were tested. The results of native listeners of Dutch were compared with those of native listeners of British English. The study investigated increased lexical activation both in the presence and in the absence of perceptual ambiguity. Thus, near-words were either based on contrasts which the nonnative listeners found difficult to distinguish or on contrasts which they found easy to distinguish.

Near-words of the first type were based on a British English phoneme contrast which has been shown to be difficult for Dutch listeners to distinguish (Broersma, 2005, submitted; Schouten, 1975), namely the vowels /æ/ and /ɛ/. As Dutch has only one vowel in the phonetic space of English /æ/ and /ɛ/, Dutch listeners are expected to initially assimilate both vowels to the same native language category (Best, 1994). Indeed, even Dutch listeners with a high level of English proficiency were found to have difficulty distinguishing between the two vowels. Although they did distinguish the vowels with a level of accuracy above chance in phonetic categorization, the Dutch listeners' accuracy was lower than that of native English listeners (Broersma, 2005, submitted).

Near-words of the second type were based on British English word-final obstruent voicing contrasts, which Dutch listeners have been shown to categorize as accurately as English listeners did (Broersma, 2005, submitted). The contrasts /z/-/s/, /v/-/f/, /b/-/p/, and /d/-/t/ were used. All of these phonemes exist in Dutch in a form very similar to the English phonemes. However, Dutch does not allow for voiced obstruents at the end of words in isolation. Thus, the voicing contrast does not occur in word-final position in Dutch. Nevertheless, Dutch listeners were found to categorize the contrasts as accurately in final position as in initial position, and as accurately as English listeners did. Further, their perception of word-final obstruents was unbiased (Broersma, 2005, submitted).

It is possible that near-words based on the consonant voicing contrast may cause increased lexical activation in an asymmetrical way. Dutch listeners might disregard the voicing of a final consonant more often when it is voiceless, which is the form they are familiar with from their native language. Further, as underlyingly voiced final obstruents are devoiced in Dutch, Dutch speakers may regularly devoice word-final obstruents in English, while hypercorrections, with a voiced pronunciation of a voiceless obstruent, may be less frequent. This may be reflected in the lexical representations of Dutch listeners. Therefore, near-words with a voiceless final consonant may cause more activation of the nearest word than vice versa. Thus, the near-word *glope* might cause more activation of *globe* than the near-word *cheab* of *cheap*.

Several experimental paradigms were used to assess the occurrence of spurious lexical activation. The most straightforward way to study lexical activation is with an auditory lexical decision experiment. In this paradigm listeners decide whether they recognize a stimulus as a word or not. Positive lexical decisions are taken to reflect a high level of lexical activation, and negative lexical decisions a low level of lexical activation. Thus, if listeners make a positive lexical decision after hearing the near-word *lemp*, this is taken to reflect a high level of activation of the near lexical neighbor *lamp*.

A task during which participants are likely to be less aware of lexical processing, but which nevertheless provides information about lexical activation, is the phonetic categorization paradigm. Listeners have often been shown to use lexical knowledge for phonetic categorization. They are inclined to classify phonemes in such a way that the input is consistent with an existing word (for a review see Pitt & Samuel, 1993). Lexical effects have been interpreted by some as evidence for feedback from the lexical to the prelexical level (e.g., McClelland & Elman, 1986). Norris, McQueen, and Cutler (2000), on the other hand, present a model in which phoneme decisions are made in decision units set up especially for this task. Decision units receive input from both the lexical and the prelexical level. Thus, lexical influences can be explained without feedback. Either way, the size of the lexical effect reflects the amount of lexical activation. In Experiment 1, the occurrence of lexical effects in phonetic categorization was investigated to assess the amount of lexical activation caused by near-words.

Experiment 1

In this experiment both an auditory lexical decision task and a phonetic categorization task were used to assess the occurrence of spurious lexical activation.

Listeners were presented with words and with near-words which they may encounter in natural speech. In the English spoken by Dutch speakers, near-words may occur as mispronunciations of the corresponding real words. In the speech of native speakers of English near-words are not likely to occur as such, but near-words may be embedded in other words or combinations of words. As lexical decisions had to be made about the speech input, however, the aim was to make the items sound natural. Excising items from a longer context would result in less naturally sounding materials, and might stimulate listeners to base their lexical decisions partially on the naturalness of the excision. In order to avoid this, all items, though spoken by a native speaker of British English, were recorded in isolation. In the phonetic categorization task, participants categorized natural tokens of the target phonemes in words and in near-words. Note that no continuum with ambiguous tokens was used (cf. Eimas, Marcovitz Hornstein, & Payton, 1990).

The hypothesis being tested was that near-words would cause more activation of their base words for nonnative listeners than for native listeners. In the lexical decision task, Dutch listeners were predicted to give more positive lexical decisions to near-words than English listeners. In the phonetic categorization task, Dutch listeners were predicted to make more errors on near-words than English listeners. If listeners made more errors on phonemes in near-words than on phonemes in real words, this would be evidence for the use of lexical knowledge for phonetic categorization. It would show that the lexical representation was not only active after presentation of the real word, but also after presentation of the near-word. The stronger the activation of the lexical representation after hearing a near-word, the more likely it would be that the categorization decision would be consistent with the real word form. Thus, the size of the lexicality effect is an indication of the activation level of the real word form after presentation of a near-word.

Method

Participants

Twenty-four native speakers of Dutch and 24 native speakers of British English took part in the lexical decision task, and 20 other native speakers of Dutch and 20 other native speakers of British English in the phonetic categorization task. The Dutch participants had a high level of proficiency in English as a second language. They had received at least 6 and on average 8 years of English instruction in primary and secondary education. The English participants did not know any Dutch. The Dutch participants were recruited from the Max Planck Institute participant pool. The English participants in the lexical decision task were recruited at the University of Birmingham, and the English participants in the phonetic categorization task from the participant pool of the Laboratory of Experimental Psychology of the University of Sussex. None reported any hearing loss. All were volunteers and received a small fee for participation.

Materials

The following five pairs of target phonemes were used: /æ/-/ɛ/, /z/-/s/, /v/-/f/, /b/-/p/, and /d/-/t/. For the lexical decision task 64 monosyllabic English words were selected as experimental items. All of them contained a target phoneme in the appropriate position. Of the first 32 words, 16 contained the vowel /ɛ/, and the other 16 /æ/. The other 32 words contained a target obstruent in word-final position, with each of the eight target obstruents appearing in final position in four of the words. Further selection criteria were that the word did not sound like an existing Dutch word, and that replacement of the target phoneme with its counterpart did not result in an existing English or Dutch word. For each experimental word, a near-word was formed by replacing the target phoneme with its counterpart (e.g., *desk* became *dask*, *globe* became *glope*). The experimental words and near-words are listed in Appendix A. With the CELEX lexical database of British English (Baayen, Piepenbrock, & Gulikers, 1995), the logarithmic lemma frequencies per million of the experimental words were determined. The experimental words containing a target vowel had a mean frequency of 2.05 and those with a target consonant a mean of 1.99. The target words were divided into two lists, which were balanced for frequency, with equal numbers of each target sound in each list. Each participant heard the 32 words from one list in their real word form and the 32 words from the other list in near-word form. Next, 68 monosyllabic English words and 68 non-words, formed by replacing either the vowel or the final consonant in a real word in such a way that the phonotactic constraints of English were not violated, were selected as filler items. 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 soundproof booth using a high quality microphone onto digital audiotape and downsampled to 16 kHz during transfer to a computer.

For the phonetic categorization task, the 64 experimental words and the 64 near-words were used.

Design

For the lexical decision task, items from the first list were presented in their real word form and those from the second list in near-word form to half of the participants, and vice versa to the other half of the participants. Each participant was presented with all of the filler words and filler non-words, so that each participant heard a total number of 100 words and 100 non- and near-words. Items were presented in a semi-random order, with the restriction that minimally three other items appeared between two target words and between two near-words.

For the phonetic categorization task the items were presented in five blocks, each block testing one phoneme contrast. The contrasts were /æ/-/ɛ/, /z/-/s/, /v/-/f/, /b/-/p/, and /d/-/t/. The vowel block contained 32 words and 32 near-words, and each consonant block contained 8 words and 8 near-words. Each block contained four repetitions of all items, semi-randomized such that the same phoneme occurred maximally five times in succession, and corresponding words and near-words were separated by at least two other items.

Procedure

Participants were tested one at a time in a quiet room. For the lexical decision task, they received written instructions in their native language, informing them that they were going to hear English words and non-words. They were asked to press a green response button, labeled “yes”, with their dominant hand if they thought the presented item was an English word, and a red response button, labeled “no” with their non-dominant 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.

For the phonetic categorization task, participants received written instructions in their native language, informing them that they would hear a series of items containing either one of two sounds. They were instructed to decide which one of two alternatives this sound was, and to indicate their response with a button press. Participants were asked to respond both as fast and as accurately as possible. Before each block, they received further information about the two response alternatives in that block, and about the position of the target phoneme. Before the /æ/-/ɛ/ block, participants heard some examples of non-words containing these phonemes while the corresponding label (“A” or “E”) appeared on a computer screen, to make it clear, particularly to the Dutch participants, which sounds were intended. The other phonemes were not expected to cause uncertainty and were not illustrated with examples. Each block started with a practice part, with 12 trials for the vowel block and 8 for the consonant blocks. The response buttons were labeled “A” and “E”, “Z” and “S”, “V” and “F”, “B” and “P”, or “D” and “T”, respectively.

Both tasks were 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. No time limit was imposed for the responses. After each button press, presentation of the next item started.

Results

In this experiment and all the following, reaction times (RTs) were measured from item offset, and outliers were removed from the analyses. Arcsine transformations were applied before analysis of proportions of correct responses or “yes” responses. Analyses are presented in tables with numbered rows. For example, the reference ‘Table 1; 2’ refers to the analysis in the second row of Table 1. The terms ‘original phoneme’ and ‘original voicing’ are used to refer to the vowel or the final consonant voicing in the real word. Thus, the original phoneme of both *lamp* and *lemp* is /æ/, and the original voicing of *globe* and *glope* is voiced. Only results directly related to the central question of this study are presented in the Results sections. Some other results are discussed in the General Discussion.

Lexical decision

Analyses of variance (ANOVAs) on proportions of “yes” responses across participants (*F1*) and across items (*F2*) are presented in Table 1. The results of one experimental pair (*ship-shib*) had to be excluded due to an error in the item lists.

The hypothesis being tested was that near-words cause more activation of the nearest word form for nonnative listeners than for native listeners. Dutch listeners were predicted to give more positive lexical decisions to near-words than English listeners. Figure 1 shows the percentage of “yes” responses, that is the percentage of items accepted as a word, for target words, filler non-words, and near-words. Note that “yes” was the correct response to words, but the incorrect response to non-words and near-words. The figure clearly shows that the Dutch and English listeners’ results differed considerably for the near-words. There was a significant interaction between native language and condition (Table 1; 1).

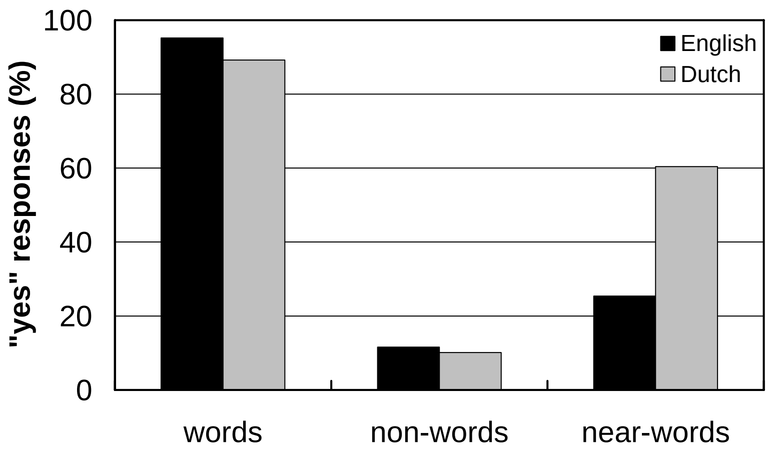

Figure 1. Experiment 1, lexical decision: English and Dutch listeners' percentage of "yes" responses to words, non-words, and near-words.

First, consider the results for the filler non-words. Filler non-words mismatched real words on phoneme contrasts which were expected to be easy to distinguish for Dutch listeners. The error rate for the filler non-words was 10.1 % for the Dutch listeners and 11.6 % for the English listeners, which was not significantly different (Table 1; 2). This result showed that the Dutch listeners did not have a bias towards "yes" responses compared to the English listeners, which is relevant for the interpretation of the following results.

In Figure 1, the results of the experimental words and near-words based on replacement of a vowel and those based on replacement of a consonant are collapsed. Below, the results are discussed in more detail for words and near-words based on a vowel contrast and those based on a consonant contrast separately.

The proportion of "yes" responses is taken to reflect the level of activation of the best-matching lexical representations. Therefore, if near-words activated lexical representations more strongly for nonnative listeners than for native listeners, this would result in an interaction between native language and condition (words vs. near-words), with a relatively large number of "yes" responses to near-words for the Dutch listeners.

Vowels

Table 2 shows the results for the words and near-words based on a vowel replacement. The table shows the percentage of real words correctly judged as a word and the percentage of corresponding near-words misjudged as a word. Results are split for the two original phonemes.

Table 1. Experiment 1, lexical decision: Analyses.

V+C	1	Language * condition	$F1$ (2, 92) = 100.36, $p < .001$	$F2$ (2, 191) = 55.25, $p < .001$
	2	Non-words; language	$F1$ (1, 47) < 1	$F2$ (1, 67) = 1.37, $p > .1$
	3	Words; language	$F1$ (1, 46) = 22.36, $p < .001$	$F2$ (1, 62) = 8.60, $p < .01$
V	4[a]	Language * condition	$F1$ (1, 46) = 64.34, $p < .001$	$F2$ (1, 30) = 42.60, $p < .001$
	5[a]	Original phoneme	$F1$ (1, 46) = 53.45, $p < .001$	$F2$ (1, 30) = 12.51, $p < .001$
	6	Words; language	$F1$ (1, 46) = 15.12, $p < .001$	$F2$ (1, 30) = 7.10, $p < .05$
	7	Near-words; language	$F1$ (1, 46) = 35.09, $p < .001$	$F2$ (1, 30) = 27.40, $p < .001$
	8	Dutch; condition	$F1$ (1, 23) = 46.46, $p < .001$	$F2$ (1, 30) = 21.18, $p < .001$
	9	English; condition	$F1$ (1, 23) = 384.18, $p < .001$	$F2$ (1, 30) = 139.49, $p < .001$
C	10	Language * condition	$F1$ (1, 46) = 45.65, $p < .001$	$F2$ (1, 30) = 40.69, $p < .001$
	11^	Original voicing	$F1$ (1, 46) = 10.87, $p < .01$	$F2$ (1, 29) = 2.79, $p > .1$
	12	Words; language	$F1$ (1, 46) = 3.18, $p = .081$	$F2$ (1, 30) = 1.90, $p > .1$
	13	Near-words; language	$F1$ (1, 46) = 74.37, $p < .001$	$F2$ (1, 30) = 56.11, $p < .001$
	14	Dutch; condition	$F1$ (1, 23) = 80.11, $p < .001$	$F2$ (1, 30) = 95.86, $p < .001$
	15	English; condition	$F1$ (1, 23) = 832.91, $p < .001$	$F2$ (1, 30) = 524.56, $p < .001$
V vs. C	16[b]	Language * condition * type	$F1$ (1, 46) < 1	$F2$ (1, 61) < 1
	17[b]	Condition * type	$F1$ (1, 46) = 38.62, $p < .001$	$F2$ (1, 61) = 21.93, $p < .001$
	18	Words; type	$F1$ (1, 46) = 12.58, $p < .001$	$F2$ (1, 61) = 4.35, $p < .05$
	19	Near-words; type	$F1$ (1, 46) = 32.09, $p < .001$	$F2$ (1, 61) = 11.54, $p < .001$

V+C: vowels and consonants; *V*: vowels; *C*: consonants; *V vs. C*: vowels versus consonants.
[a, b]: Results with the same superscript were found with a single analysis.
^: This factor was not included in the other analyses.

Table 2. Experiment 1, lexical decision: English and Dutch listeners' percentage of "yes" responses to words and near-words, separately for items with an /æ/, /ɛ/, voiced (+V) or voiceless final consonant (-V). Examples are given in brackets.

	English	**Dutch**
word, /æ/ *(lamp)*	92.2	77.1
word, /ɛ/ *(desk)*	95.8	93.2
near-word, /æ/→/ɛ/ *(lemp)*	20.9	53.6
near-word, /ɛ/→/æ/ *(dask)*	56.8	77.9
word, +V *(globe)*	95.3	89.6
word, -V *(cheap)*	97.8	97.8
near-word, +V→-V *(glope)*	9.4	55.0
near-word, -V→+V *(cheab)*	13.9	55.0

As predicted, there was a significant interaction between native language and condition (Table 1; 4). The Dutch listeners gave less "yes" responses to real words than the English listeners did (Table 1; 6), but they gave more "yes" responses to near-words than the English listeners did (Table 1; 7). The proportion of "yes" responses was higher for words than for near-words, both for the Dutch listeners (Table 1; 8) and for the English listeners (Table 1; 9).

Because of the large proportion of incorrect responses, the RTs of the correct responses were not analyzed.

Table 2 shows that there were more “yes” responses to words containing an /ɛ/ than to words containing an /æ/, and the same for the corresponding near-words. This is discussed in detail in the General Discussion.

Consonants

Table 2 also shows the results for the words and near-words based on a consonant replacement. It shows the percentage of real words correctly judged as a word and the percentage of corresponding near-words misjudged as a word. Results are split for voiced and voiceless original phonemes.

As predicted, there was a significant interaction between native language and condition again (Table 1; 10). Whereas there was no difference between the Dutch and the English listeners’ proportion of “yes” responses to the words (Table 1; 12), the Dutch listeners gave more “yes” responses to near-words than the English listeners did (Table 1; 13). The proportion of “yes” responses was higher for words than for near-words, both for the Dutch listeners (Table 1; 14) and for the English listeners (Table 1; 15). Because of the large proportion of incorrect responses, the RTs of the correct responses were not analyzed.

Vowels versus consonants

There was no three-way interaction among native language, condition, and type (vowels vs. consonants) (Table 1; 16).

Phonetic categorization

Analyses were carried out on proportions of correct responses. If listeners based their categorization decision on the nearest lexical representation, this would result in a correct response for real words and an incorrect response for near-words. Therefore, if near-words activated lexical representations more strongly for nonnative listeners than for native listeners, this would result in an interaction between native language and condition (words vs. near-words), with a relatively large number of incorrect responses to near-words for the Dutch listeners. Figure 2 shows that this is exactly the pattern found in the results. Indeed, there was a significant interaction between native language and condition (Table 3; 1). In Figure 2, the results of the words and near-words based on replacement of a vowel and those based on replacement of a consonant are collapsed. Below, the results are discussed in more detail for words and near-words based on a vowel contrast and those based on a consonant contrast separately.

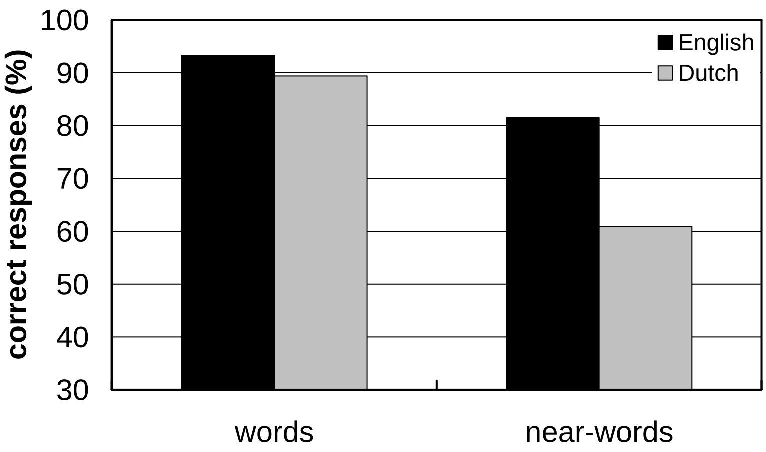

Figure 2. Experiment 1, phonetic categorization: English and Dutch listeners' percentage of correct responses to words and near-words.

Table 3. Experiment 1, phonetic categorization: Analyses.

V+C	1	Language * condition	$F1$ (1, 38) = 20.98, $p < .001$	$F2$ (1, 63) = 23.36, $p < .001$
V	2[a]	Language * condition	$F1$ (1, 38) = 22.69, $p < .001$	$F2$ (1, 30) = 26.45, $p < .001$
	3[a]	Language	$F1$ (1, 38) = 151.94, $p < .001$	$F2$ (1, 30) = 143.58, $p < .001$
	4[a]	Original phoneme * condition	$F1$ (1, 38) = 98.00, $p < .001$	$F2$ (1, 30) = 129.88, $p < .001$
	5	Dutch; condition	$F1$ (1, 19) = 107.93, $p < .001$	$F2$ (1, 30) = 320.92, $p < .001$
	6	English; condition	$F1$ (1, 19) = 123.95, $p < .001$	$F2$ (1, 30) = 62.25, $p < .001$
	7	Dutch; > 50%	t (19) = 7.00, $p < .001$	t (31) = 7.28, $p < .001$
	8	Words; language	$F1$ (1, 38) = 21.92, $p < .001$	$F2$ (1, 30) = 37.24, $p < .001$
C	9	Language * condition	$F1$ (1, 38) = 5.33, $p < .05$	$F2$ (1, 31) = 4.52, $p < .05$
	10^	Original voicing	$F1$ (1, 38) = <1	$F2$ (1, 30) = 2.39, $p > .1$
	11	Dutch; condition	$F1$ (1, 19) = 36.13, $p < .001$	$F2$ (1, 31) = 19.18, $p < .001$
	12	English; condition	$F1$ (1, 19) = 6.92, $p < .05$	$F2$ (1, 31) = 6.15, $p < .05$
	13	Dutch; > 50%	t (19) = 17.60, $p < .001$	t (31) = 26.44, $p < .001$
	14	Words; language	$F1$ (1, 38) < 1	$F2$ (1, 31) < 1
V vs.	15[b]	Language * condition * type	$F1$ (1, 38) = 4.63, $p < .05$	$F2$ (1, 62) = 5.91, $p < .05$
C	16[b]	Condition * type	$F1$ (1, 38) = 91.44, $p < .001$	$F2$ (1, 62) = 18.77, $p < .001$
	17	Words; type	$F1$ (1, 38) = 18.01, $p < .001$	$F2$ (1, 62) = 10.13, $p < .01$
	18	Near-words; type	$F1$ (1, 38) = 181.62, $p < .001$	$F2$ (1, 62) = 73.31, $p < .001$

V+C: vowels and consonants; *V*: vowels; *C*: consonants; *V vs. C*: vowels versus consonants.
[a, b]: Results with the same superscript were found with a single analysis.
^: This factor was not included in the other analyses.

Table 4. Experiment 1, phonetic categorization: English and Dutch listeners' percentage of correct responses to words and near-words, separately for items with an /æ/, /ɛ/, voiced (+V) or voiceless final consonant (-V). Examples are given in brackets.

	English	Dutch
word, /æ/ *(lamp)*	89.3	78.9
word, /ɛ/ *(desk)*	96.8	90.5
near-word, /æ/→/ɛ/ *(lemp)*	93.4	50.9
near-word, /ɛ/→/æ/ *(dask)*	49.8	20.5
word, +V *(globe)*	90.5	93.5
word, -V *(cheap)*	96.6	94.8
near-word, +V→-V *(glope)*	91.2	84.2
near-word, -V→+V *(cheab)*	90.0	88.2

Vowels

Table 4 shows the percentage of correct responses for the vowel targets, separately for words and near-words and for the two original phonemes.

As predicted, there was a significant interaction between native language and condition (Table 3; 2). The proportion of correct responses was higher for real words than for near-words, both for the Dutch listeners (Table 3; 5) and for the English listeners (Table 3; 6). However, the interaction between native language and condition showed that this lexicality effect was stronger for the Dutch listeners than for the English listeners. Thus, near-words activated the nearest lexical representation more for nonnative listeners than for native listeners.

Table 4 shows that there were more correct responses to items containing an /ɛ/ than to items containing an /æ/. This is further discussed in the General Discussion.

Consonants

Table 4 also shows the percentage of correct responses for the consonant targets, separately for words and near-words and for voiced and voiceless original phonemes.

As predicted, there was an interaction between native language and condition (Table 3; 9). The proportion of correct responses was significantly higher for real words than for near-words, both for the Dutch listeners (Table 3; 11) and for the English listeners (Table 3; 12). However, the interaction showed that near-words activated lexical representations more strongly for Dutch listeners than for English listeners. Thus, there was more activation of spurious lexical competitors for the nonnative listeners than for the native listeners.

Vowels versus consonants

There was a significant three-way interaction among native language, condition, and type (Table 3; 15). There was more spurious lexical activation for Dutch listeners than for English listeners, both for vowel items and for consonant items, but the difference was larger for the vowel items than for the consonant items.

Discussion

Both tasks showed that near-words caused more lexical activation for the Dutch listeners than for the English listeners. In the lexical decision task, the Dutch listeners accepted near-words as words more often than the English listeners did. The discrepancy between near-words and real words affected the Dutch listeners' responses less than the English listeners' responses. Presentation of a near-word resulted in more activation of the nearest lexical representation for the Dutch listeners than for the English listeners.

In the phonetic categorization task, the effect of lexical status, with a lower proportion of correct responses to near-words than to real words, was larger for the Dutch listeners than for the English listeners. Thus, categorization responses to a near-word were more often consistent with the nearest word form for nonnative listeners than for native listeners. This strong lexical influence on the categorization of near-words indicated a high level of activation of the real word form upon presentation of the near-word for nonnative listeners.

The results were similar for vowel items and consonant items. In the lexical decision task, there was no difference between vowel and consonant items. In the phonetic categorization task, the activation increase for the Dutch listeners (compared to the English listeners) was larger for the vowel items than for the consonant items, but for both types of stimuli near-words caused more lexical activation for the Dutch listeners than for the English listeners.

For the items which involved replacement of /æ/ with /ɛ/ or vice versa, this may have resulted from the Dutch listeners' inability to distinguish between the two vowels. As pointed out in the Introduction, Broersma (2005, submitted) found that Dutch listeners (meeting the same description as the Dutch listeners in the present experiment) distinguished /æ/ and /ɛ/ less accurately than English listeners did. It is likely that the vowels /æ/ and /ɛ/ were ambiguous for the Dutch listeners in the present experiment as well. As a near-word with a perceptually ambiguous phoneme may not mismatch with the nearest lexical representation, this lexical representation might remain highly activated after perception of the ambiguous phoneme.

The phoneme pairs that were used for the consonant items should not lead to perceptual ambiguity for the Dutch listeners, as Dutch listeners from the same population have been demonstrated to distinguish between English voiced and voiceless final obstruents with a native-like level of accuracy (Broersma, 2005, submitted). Therefore, the present results for the near-words with a consonant substitute provide evidence that spurious lexical activation is not restricted to words containing perceptually ambiguous phonemes.

Thus, Experiment 1 showed that near-words activated the nearest word form more strongly for nonnative listeners than for native listeners. For a further understanding of the role of near-words in the comprehension of nonnative speech, it is important to gain insight in the amount of spurious activation that near-words cause. The higher the level of activation of the spurious lexical competitor, the more it complicates nonnative speech comprehension. Therefore, in Experiment 2, a cross-modal priming task was used to investigate to what extent words and near-words activate the corresponding lexical representations. The cross-modal priming task allows for the measurement of the amount of activation of a visually presented word caused by a preceding auditory prime. Additionally, a phonetic categorization task was used.

Experiment 2

This experiment consisted of two tasks, a cross-modal priming task and a phonetic categorization task.

In the cross-modal priming task, two questions were examined. First, the task aimed to provide further support for the results from Experiment 1 that near-words caused more lexical activation for nonnative listeners than for native listeners. To this end, the results of native and nonnative listeners were compared. For both groups of listeners, facilitation was expected when the visual target word was preceded by the auditory presentation of the identical real word. For the English listeners, near-words were expected to lead to less lexical activation than real words. Therefore, after near-words, less facilitation was expected than after real words, or possibly no facilitation at all (cf. Marslen-Wilson, Nix, & Gaskell, 1995). For the Dutch listeners, the discrepancy between the near-word and the word might not prevent or diminish the activation of the lexical representation as much as for the English listeners. Thus, for the Dutch listeners, there might be more facilitation after a near-word than for the English listeners. This would provide further evidence that near-words caused more spurious lexical activation for nonnative listeners than for native listeners.

Second, this task assessed the level of lexical activation caused by a near-word compared to the level of activation caused by a real word for the nonnative listeners. If a mismatch between a near-word and a lexical representation led to a decrease in the level of lexical activation, there would be less facilitation after a near-word than after a real word. On the other hand, if for the nonnative listeners near-words and lexical representations did not mismatch, near-words would activate lexical representations to the same extent as real words did, and there should be as much facilitation after a near-word as after a real word.

The phonetic categorization task also aimed to find further evidence for the finding of Experiment 1 that near-words activated lexical representations more for nonnative listeners than for native listeners. As in Experiment 1, if in a phonetic categorization task listeners made more errors on phonemes in near-words than on phonemes in real words, this would be evidence for lexical activation after presentation of the near-word.

Near-words are not likely to be encountered in isolation in the English of native speakers, but they are likely to occur embedded in other words. As a lexical decision on the auditory stimulus was required in Experiment 1, the materials of that experiment were recorded in isolation. In the present experiment, however, lexical decisions were not made on the auditory stimuli. Therefore, words and corresponding near-words were recorded in a carrier context, in which they occurred as an initial embedding. For example, *DEFinite* and *DAFfodil* were recorded for the word and near-word pair *deaf-daf*. As disambiguating information would suppress the activation of the embedded word (Zwitserlood, 1989), any previous activation of the embedded word would not be observable after presentation of the whole carrier word. Therefore, words and near-words were excised from their carrier contexts to serve as auditory primes (e.g., *deaf* was excised from *definite*, and *daf* from *daffodil*).

Method

Participants

Thirty-six native speakers of Dutch and 36 native speakers of British English, meeting the description given in Experiment 1 took part. None of them 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. None of the participants reported any visual loss or reading disability.

Materials

No appropriate set of items based on the final consonant voicing contrast could be found. Therefore, only items based on the /æ/-/ɛ/ contrast were used.

For the cross-modal priming task, twelve monosyllabic English words were selected as visual target words, six of which contained an /æ/ and six an /ɛ/. The mean logarithmic lemma frequency per million of the visual target words, determined with the CELEX lexical database (Baayen et al., 1995), 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 and 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 Identity condition (e.g., auditory presentation of *deaf* from *definite*, visual presentation of *deaf*). The almost embedded 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*). The resulting auditory primes were real words in the Identity condition, near-words in the Mismatch condition, and non-words in the Control condition. The experimental targets, primes and carriers are listed in Appendix B.

Twenty-eight filler words and 32 filler non-words with identical primes, and the same number of words and non-words with mismatching primes and with unrelated primes were selected and constructed as described for the experimental items. Mismatching primes differed from the visual items in one phoneme. The mismatch concerned a vowel for 12 and a consonant for 16 of the filler words, and 16 vowels and 16 consonants for the filler non-words. Items selected for visual presentation were not spelled like existing Dutch words, and items selected for auditory presentation did not sound like existing Dutch words.

For the phonetic categorization task, the 12 pairs of words and near-words which served as experimental Identity and Mismatch primes were used.

Design

In the cross-modal priming task, each participant was presented with each of the experimental visual targets only once, with four of the targets in each of the three conditions: Identity condition (preceded by an identical auditory prime), Mismatch condition (preceded by a prime which mismatched the target in the /æ/-/ɛ/ contrast), and Control condition (preceded by a phonologically and semantically unrelated prime). Each participant was presented with all of the filler words and filler non-words, so that each participant saw a total of 96 words and 96 non-words, with 64 presentations in each of the three conditions. While the experimental Mismatch condition only involved vowel replacements, the total stimulus set contained an equal number of vowel and consonant replacements. Items were presented in a semi-random order, such that maximally five visually presented words or five visually presented non-words occurred in succession, and two experimental targets were separated by at least one other item.

In the phonetic categorization task, all items were presented four times, in a semi-random order as in Experiment 1.

Procedure

Participants were tested one at a time in a quiet room. All participants did both the cross-modal priming task and the phonetic categorization task, with a short break in between.

First, for the cross-modal priming task, the participants received written instructions in their native language, informing them that on each trial they would hear part of an English word, directly after which an English word or non-word would appear on a computer screen. They were asked to press a green response button, labeled "yes", with their dominant hand if they thought the visually presented item was an English word, and a red response button, labeled "no", with their non-dominant hand if they thought the visually presented item was not an English word. Participants were asked to respond both as fast and as accurately as possible. The experiment started with 12 practice trials and was controlled with NESU software. On each trial, an auditory stimulus was presented and at offset of that, a visual stimulus was presented. The auditory materials were presented binaurally over closed headphones at a comfortable listening level and the visual materials appeared in large font on a computer screen in front of the participants. Participants responded by pressing one of the two response buttons. No time limit was imposed for the responses. After each button press, the next trial started.

After having finished the cross-modal priming task, participants did the phonetic categorization task. The procedure was as described for Experiment 1, except that the present task consisted of one block only, testing the /æ/-/ɛ/ contrast.

Results

Cross-modal priming

For the experimental items, visual targets were always real words and the correct response was "yes". In the following ANOVAs, the dependent variable was the mean RT of the correct responses. Two main questions were examined. First, the results of the native and nonnative listeners were compared. For both groups of listeners, facilitation was expected in the Identity condition. Thus, RTs were expected to be shorter in the Identity condition than in the Control condition. If near-words caused more spurious lexical activation for nonnative listeners than for native listeners, the Dutch and English listeners' results would differ in the Mismatch condition. For the English listeners, there would be less facilitation in the Mismatch condition than in the Identity condition or possibly no facilitation at all. For the Dutch listeners, there would be more facilitation in the Mismatch condition than for the English listeners. Figure 3 shows that the pattern of results was exactly as expected. Second, for the Dutch listeners, the level of lexical activation caused by near-words was compared to the activation caused by real words. Thus, the amount of facilitation in the Mismatch condition was compared to that in the Identity condition.

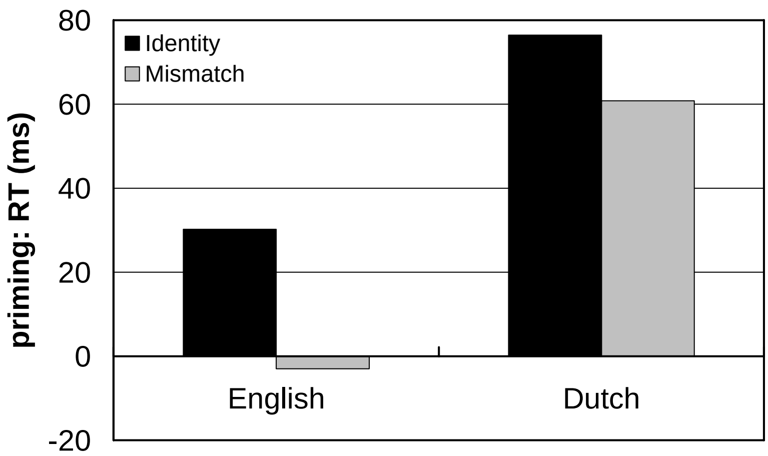

Figure 3. Experiment 2, cross-modal priming: English and Dutch listeners' priming results, computed as the difference between the reaction times of correct responses in the Identity or the Mismatch condition and the Control condition, with a positive value indicating facilitation.

Table 5. Experiment 2, cross-modal priming: Analyses.

RT	1[a]	Language * condition	*F1* (2, 140) = 4.74, $p < .01$	*F2* (2, 22) = 3.35, $p = .054$
	2[a]	Language	*F1* (1, 70) = 40.03, $p < .001$	*F2* (1, 11) = 39.95, $p < .001$
	3	Dutch; condition	*F1* (2, 70) = 7.45, $p < .001$	*F2* (2, 22) = 4.77, $p < .05$
	4	English; condition	*F1* (2, 70) = 7.65, $p < .001$	*F2* (2, 22) = 5.18, $p < .05$
	5	Dutch; Identity vs. Control	*F1* (1, 35) = 11.04, $p < .01$	*F2* (1, 11) = 8.38, $p < .05$
	6	Dutch; Mismatch vs. Control	*F1* (1, 35) = 8.89, $p < .01$	*F2* (1, 11) = 4.86, $p < .05$
	7	Dutch; Identity vs. Mismatch	*F1* (1, 35) < 1	*F2* (1, 11) < 1
	8	English; Identity vs. Control	*F1* (1, 35) = 13.58, $p < .01$	*F2* (1, 11) = 8.02, $p < .05$
	9	English; Mismatch vs. Control	*F1* (1, 35) < 1	*F2* (1, 11) < 1
	10	English; Identity vs. Mismatch	*F1* (1, 35) = 11.67, $p < .01$	*F2* (1, 11) = 10.56, $p < .01$
	11^	Phoneme	*F1* (1, 57) = 24.47, $p < .001$	*F2* (1, 10) = 2.75, $p > .1$
Corr	12[b]	Phoneme	*F1* (1, 67) = 47.54, $p < .001$	*F2* (1, 10) = 3.74, $p = .082$
	13[b]	Language	*F1* (1, 67) = 56.64, $p < .001$	*F2* (1, 10) = 4.69, $p = .056$

RT: reaction time; *Corr*: proportion correct
[a, b]: Results with the same superscript were found with a single analysis.
^: This factor was not included in the other analyses.

Table 6. Experiment 2, cross-modal priming: English and Dutch listeners' percentage of correct responses and RTs of correct responses for target words in Control, Identity, and Mismatch condition, separately for target words with an /æ/ or /ɛ/. Examples are given in brackets.

		correct (%)		RT (ms)	
target word	**condition (prime)**	**English**	**Dutch**	**English**	**Dutch**
/æ/ *(cat)*	**Control *(pock)***	95.7	78.1	581	733
	Identity *(cat)*	92.6	68.1	546	659
	Mismatch *(ket)*	95.3	63.8	594	683
/ɛ/ *(deaf)*	**Control *(hov)***	98.6	94.0	549	678
	Identity *(deaf)*	100	90.0	524	601
	Mismatch *(daf)*	95.7	94.4	544	614

Analyses are presented in Table 5. Table 6 shows the English and Dutch listeners' percentage of correct responses and the RTs for the correct responses in the three conditions, separately for the two target phonemes. There were no significant effects in the proportions of correct responses.

For the RTs, the interaction between native language and condition was significant in the analysis by subjects but narrowly missed significance in the analysis by items (Table 5; 1). There was a significant main effect of condition, both for the Dutch (Table 5; 3) and for the English listeners (Table 5; 4). For the Dutch listeners, RTs were shorter in the Identity condition than in the Control condition (Table 5; 5), RTs were shorter in the Mismatch condition than in the Control condition (Table 5; 6), and there was no difference between RTs in the Identity and the Mismatch condition (Table 5; 7). For the English listeners, RTs were also shorter in the Identity condition than in the Control condition (Table 5; 8). For the English listeners, however, there was no difference between RTs in the Mismatch and the Control condition (Table 5; 9), and RTs were shorter in the Identity condition than in the Mismatch condition (Table 5; 10).

Phonetic categorization

If listeners based their categorization decision on the nearest lexical representation, this would result in a correct response for real words and an incorrect response for near-words. Therefore, if near-words activated lexical representations more strongly for nonnative listeners than for native listeners, this would result in an interaction between native language and condition (words vs. near-words), with a relatively large number of incorrect responses to near-words for the Dutch listeners.

Figure 4 shows the percentage of correct responses for words and near-words. The pattern of results is as predicted, with a relatively high number of errors to near-words for the Dutch listeners. Analyses are presented in Table 7. Table 8 shows the data separately for the two original phonemes.

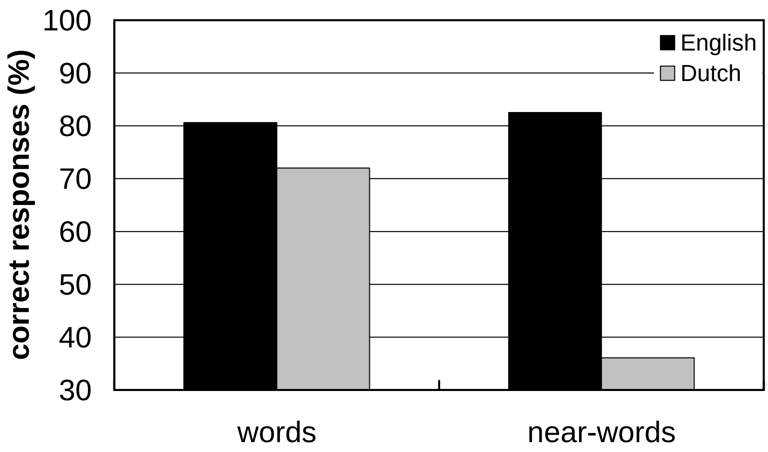

Figure 4. Experiment 2, phonetic categorization: English and Dutch listeners' percentage of correct responses to words and near-words.

Table 7. Experiment 2, phonetic categorization: Analyses.

1[a]	Language * condition	*F1* (1, 70) = 59.19, *p* < .001	*F2* (1, 10) = 22.14, *p* < .001
2[a]	Language	*F1* (1, 70) = 179.04, *p* < .001	*F2* (1, 10) = 112.73, *p* < .001
3[a]	Original phoneme * condition	*F1* (1, 70) = 184.46, *p* < .001	*F2* (1, 10) = 22.32, *p* < .001
4	Dutch; condition	*F1* (1, 35) = 86.39, *p* < .001	*F2* (1, 10) = 14.53, *p* < .01
5	English; condition	*F1* (1, 35) = 1.50, *p* > .1	*F2* (1, 10) < 1
6	Dutch; > 50%	*t* (35) = 4.48, *p* < .001	*t* (11) = 4.90, *p* < .001
7	Words; language	*F1* (1, 70) = 20.24, *p* < .001	*F2* (1, 10) = 8.82, *p* < .05

[a]: Results with the same superscript were found with a single analysis.

Table 8. Experiment 2, phonetic categorization: English and Dutch listeners' percentage of correct responses to words and near-words, separately for items with an /æ/ or /ɛ/. Examples are given in brackets.

	English	**Dutch**
word, /æ/ *(cat)*	64.1	56.8
word, /ɛ/ *(deaf)*	97.1	87.2
near-word, /æ/→/ɛ/ *(ket)*	96.3	52.0
near-word, /ɛ/→/æ/ *(daf)*	68.6	20.1

As predicted, there was a significant interaction between native language and condition (Table 7; 1). For the Dutch listeners, the proportion correct was higher for real words than for near-words (Table 7; 4). For the English listeners, there was no difference between the conditions (Table 7; 5). Thus, an effect of lexicality was only present for the Dutch listeners and not for the English listeners. This finding provided further evidence that near-words activated spurious lexical competitors more for nonnative listeners than for native listeners.

Discussion

First, these results support those of Experiment 1. The results of both tasks showed that near-words caused more activation of lexical representations for nonnative listeners than for native listeners. In the cross-modal priming task, for both groups of listeners, recognition of the visual target word was facilitated when it was preceded by auditory presentation of the same word. For the English listeners, recognition of the target word was not facilitated when it was preceded by the corresponding near-word. For the Dutch listeners on the other hand, recognition of the visual target word was also facilitated when it was preceded by the corresponding near-word. This shows that the lexical representation was active after presentation of the near-word, and that the discrepancy between the near-word and the word did not lead to (full) deactivation of the lexical representation. In the phonetic categorization task, for native listeners, there was no evidence that the lexical representations were active after presentation of the near-words. For the nonnative listeners, the proportion of correct responses was lower for near-words than for words. This showed that lexical representations did influence the categorization responses to near-words, and thus that lexical representations were active after presentation of near-words.

Second, the results of the cross-modal priming task provided further insight in the amount of lexical activation caused by near-words. For the Dutch listeners, there was as much facilitation after a near-word as after a real word. Auditory stimuli like *deaf* and *daf* activated the word form (*deaf*) to the same extent. The presentation of the vowel in a near-word did not lead to bottom-up inhibition but to a further activation of the lexical representation, similar to the presentation of the original phoneme in a real word. This suggests that the influence of near-words on nonnative speech comprehension may be considerable. The level of lexical activation caused by a near-word is very high (as high as lexical activation caused by the real word), and spurious lexical activation may not be easy to overcome. Therefore, spurious lexical competition may greatly complicate nonnative speech comprehension.

The cross-modal priming task provided detailed information about the amount of lexical activation caused by near-words and words. In the present experiment, only items based on the replacement of the phonemes /æ/ and /ɛ/ could be used. In order to gain more insight in the role of near-words based on unambiguous phonemes, another cross-modal priming experiment was carried out with both near-words based on vowel replacement and near-words based on consonant replacement.

It was argued in the Introduction that for Dutch listeners the near-word *glope* may cause more activation of *globe* than the near-word *cheab* of *cheap*. In Experiment 1, no asymmetry was found between voiced and voiceless consonants (i.e., there were no main effects of or interactions involving original voicing). However, the following cross-modal priming experiment provided an opportunity to measure small differences in lexical activation of near-words with voiced and voiceless consonants on-line. There were no a-priori reasons to expect an asymmetry in the occurrence of spurious lexical activation for the near-words based on the /æ/-/ɛ/ contrast. Indeed, in the cross-modal priming task of Experiment 2, no such asymmetry was found (i.e., there were no interactions involving original voicing).

Experiment 3

This experiment consisted of a cross-modal priming task and a phonetic categorization task.

The previous experiments have provided evidence that near-words caused more spurious lexical activation for nonnative listeners than for native listeners. Both tasks of the present experiment aimed to provide further evidence for these findings with items which were constructed in a different manner. In Experiment 1, auditory stimuli were recorded in isolation. In Experiment 2, words and near-words were word-initially embedded in a single carrier word. In the present experiment, words and near-words were recorded embedded in two words, spanning a word boundary. The words and near-words from Experiment 1 were used. For the full set of items, an appropriate two-word carrier context was found. Thus, word and near-word pairs like *lamp* and *lemp* were excised from the carrier contexts *eviL AMPlitude* and *eviL EMPire*, and *globe* and *glope* were excised from *biG LOBE* and *biG LOPE*.

Method

Participants
Thirty-six native speakers of Dutch and 36 native speakers of British English, none of whom had participated in the previous experiments, took part in the cross-modal priming task, and 20 other native speakers of Dutch and 20 other native speakers of British English took part in the phonetic categorization task. Participants met the description given in Experiment 2.

Materials
For the cross-modal priming task, the 64 real words from Experiment 1 were used as visual targets. The same words served as auditory primes for the Identity condition (e.g., prime: *lamp*, target: *lamp*). The 64 near-words from Experiment 1 were used as auditory primes for the Mismatch condition (e.g., prime: *lemp*, target: *lamp*). Additionally, 64 monosyllabic items which were phonologically and semantically unrelated to the target word, half of which were words and half of which non-words, were selected as auditory primes for the Control condition (e.g., prime: *bike*, target: *lamp*).

For all Identity, Mismatch, and Control primes, a fragment was found in which the item occurred as an embedding (e.g., *evil amplitude* for *lamp*, *evil empire* for *lemp*, *prefab icon* for *bike*). Each carrier fragment consisted of two words, that together contained the auditory prime across the word boundary. The first one or two consonants of the auditory prime were part of the first word, and the rest of the auditory prime was part of the second word. In approximately one third of the cases, the offset of the auditory prime coincided with the offset of the second carrier word. The carrier fragments were chosen such that the phoneme environments of the Identity and the Mismatch primes were as similar as possible. For six of the experimental pairs, no suitable carrier fragments could be found. For these pairs, fictitious geographic names were used (for both the Identity and the Mismatch prime within one pair). The experimental targets, primes, and carriers are listed in Appendix A.

Next, 66 monosyllabic filler words were selected as visual targets, 22 with an Identity prime, 22 with a Mismatch prime, and 22 with an unrelated Control prime. As the Identity primes were always words, and the Mismatch primes non-words, half of the Control primes were words, the other half non-words. Finally, 129 monosyllabic filler non-words were chosen as visual targets, 43 with an Identity prime, 43 with a Mismatch prime, and 43 with a Control prime. For the filler non-words the Identity primes were always non-words, the Mismatch primes were words, and half of the Control primes were words, the other half non-words. For all fillers, half of the Mismatch primes differed from the corresponding targets in the vowel, and the other half in the final consonant. For all primes for the filler word and non-word targets, a two-word fragment was selected in which the item 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 onto a computer at a sample rate of 16 kHz. Auditory primes were extracted from the carrier fragments.

For the phonetic categorization task, the 64 pairs of words and near-words which served as experimental Identity and Mismatch primes were used.

Design

In the cross-modal priming task, each participant was presented with each of the 64 experimental targets only once. The experimental targets were assigned to the Identity, Mismatch, and Control conditions so that two conditions contained 21 experimental targets and the third 22, counterbalancing over participants. Each participant was presented with 65 of the filler words, 21 in two conditions and 22 in the third condition, so that the total number of (experimental and filler) words was 43 in each condition. Each participant was also presented with 43 filler non-word targets in each condition. The total number of trials was 258. Items were presented in a semi-random order as described for Experiment 2.

For the phonetic categorization task, the design was as described for Experiment 1.

Procedure
The procedure of the cross-modal priming task was as described for Experiment 2, and the procedure of the phonetic categorization task was as described for Experiment 1.

Results

Cross-modal priming
For the experimental items, visual targets were always real words and the correct response was “yes”. Analyses of the proportion of correct responses and of the RTs of the correct responses were used to investigate the occurrence of spurious lexical activation. In these analyses, the results of native and nonnative listeners were compared. For both groups of listeners, facilitation was expected in the Identity condition. Thus, the proportion of correct responses should be higher and/or the RTs of the correct responses should be shorter in the Identity condition than in the Control condition. If near-words caused more spurious lexical activation for nonnative listeners than for native listeners, there would be more facilitation for Dutch listeners than for English listeners in the Mismatch condition (as compared to the Control condition). Figure 5 shows exactly this pattern. In the RTs, there was a significant interaction between native language and condition (Table 9; 1). In Figure 5, the results of the items based on a vowel contrast and those based on a consonant contrast are collapsed. Below, the results are discussed separately for vowel items and consonant items in more detail.

Further, the occurrence of an asymmetry between targets with an /æ/ or an /ɛ/ and between targets with voiced and voiceless final consonants was investigated. The expectation was that near-words with a voiceless final consonant (e.g., *glope*) would cause more facilitation of real words than near-words with a voiced final consonant (e.g., *cheab*).

Vowels
Table 10 shows the English and Dutch listeners’ percentage of correct responses and the RTs for the correct responses in the three conditions, separately for the two target phonemes. There were no interactions involving condition nor a main effect of condition in the proportions of correct responses.

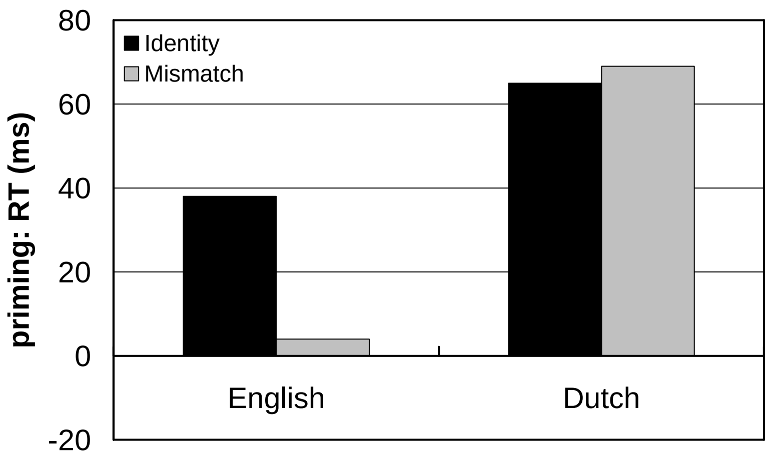

Figure 5. Experiment 3, cross-modal priming: English and Dutch listeners' priming results, computed as the difference between the reaction times of correct responses in the Identity or the Mismatch condition and the Control condition, with a positive value indicating facilitation.

In the analysis of the RTs, there was a significant interaction between native language and condition (Table 9; 2). This interaction was further investigated with ANOVAs comparing the two groups of listeners and two conditions at a time. For the Identity versus the Control condition, there was no interaction between condition and native language (Table 9; 5), and RTs were shorter in the Identity condition than in the Control condition (Table 9; 6). For the Mismatch versus the Control condition, there was a significant interaction between condition and native language (Table 9; 7). For the Dutch listeners, RTs were shorter in the Mismatch condition than in the Control condition (Table 9; 8). For the English listeners, there was no significant difference between the Mismatch condition and the Control condition (Table 9; 9). For the Identity condition versus the Mismatch condition, there was no interaction between condition and language (Table 9; 10), and there was no main effect of condition (Table 9; 11). Thus, for both groups of listeners recognition of the visual target word was facilitated when it was preceded by auditory presentation of the same word. For the Dutch listeners, recognition of the visual target word was facilitated when it was preceded by the corresponding near-word as well. For the English listeners, there was no facilitation after presentation of a near-word.

There were no interactions involving phoneme, neither in the proportions correct nor in the RTs. Thus, near-words with an /æ/ and those with an /ε/ caused a similar amount of lexical activation.

Table 9. Experiment 3, cross-modal priming: Analyses.

V+C RT	1	Language * condition	*F1* (2, 138) = 13.67, *p* < .001	*F2* (2, 126) = 9.45, *p* < .001
V RT	2[a]	Language * condition	*F1* (2, 138) = 5.10, *p* < .01	*F2* (2, 60) = 3.27, *p* < .05
	3[a]	Phoneme	*F1* (1, 69) = 21.05, *p* < .001	*F2* (1, 30) = 6.38, *p* < .05
	4[a]	Language	*F1* (1, 69) = 10.70, *p* < .01	*F2* (1, 30) = 54.11, *p* < .001
	5[b]	Identity vs. Control; language * condition	*F1* (1, 69) = 1.73, *p* > .1	*F2* (1, 30) = 1.81, *p* > .1
	6[b]	Condition	*F1* (1, 69) = 19.24, *p* < .001	*F2* (1, 30) = 18.13, *p* < .001
	7	Mismatch vs. Control; language * condition	*F1* (1, 69) = 9.59, *p* < .01	*F2* (1, 30) = 7.82, *p* < .01
	8	Dutch; Mismatch vs. Control	*F1* (1, 34) = 19.27, *p* < .001	*F2* (1, 30) = 12.61, *p* < .001
	9	English; Mismatch vs. Control	*F1* (1, 35) = 3.21, *p* = .082	*F2* (1, 30) = 3.36, *p* = .077
	10[c]	Identity vs. Mismatch; language * condition	*F1* (1, 70) = 3.86, *p* = .080	*F2* (1, 30) = 1.22, *p* > .1
	11[c]	Condition	*F1* (1, 70) < 1	*F2* (1, 30) < 1
V Corr	12[d]	Phoneme	*F1* (1, 70) = 84.05, *p* < .001	*F2* (1, 30) = 5.33, *p* < .05
	13[d]	Language	*F1* (1, 70) = 62.50, *p* < .001	*F2* (1, 30) = 31.76, *p* < .001
C Corr	14[e]	Language * condition	*F1* (2, 140) = 3.16, *p* < .05	*F2* (2, 60) = 5.18, *p* < .01
	15[e]	Voicing	*F1* (1, 70) < 1	*F2* (1, 30) < 1
	16[e]	Language	*F1* (1, 70) = 38.65, *p* < .001	*F2* (1, 30) = 29.74, *p* < .001
	17	Dutch; condition	*F1* (2, 70) = 3.69, *p* < .05	*F2* (2, 60) = 6.65, *p* < .01
	18	English; condition	*F1* (2, 70) < 1	*F2* (2, 60) = 1.22, *p* > .1
	19	Dutch; Identity vs. Control	*F1* (1, 35) = 4.02, *p* < .053	*F2* (1, 30) = 7.97, *p* < .01
	20	Dutch; Mismatch vs. Control	*F1* (1, 35) = 5.20, *p* < .05	*F2* (1, 30) = 10.46, *p* < .01
	21	Dutch; Identity vs. Mismatch	*F1* (1, 35) < 1	*F2* (1, 30) < 1
C RT	22[f]	Language * condition * voicing	*F1* (2, 138) = 7.36, *p* < .001	*F2* (2, 60) = 3.17, *p* < .05
	23[f]	Language	*F1* (1, 69) = 10.57, *p* < .01	*F2* (1, 30) = 47.11, *p* < .001
	24	Voiced; language * condition	*F1* (2, 138) = 12.30, *p* < .001	*F2* (2, 30) = 10.37, *p* < .001
	25[g]	Voiced, Identity vs. Control; language * condition	*F1* (1, 69) < 1	*F2* (1, 15) < 1
	26[g]	Condition	*F1* (1, 69) = 11.29, *p* < .001	*F2* (1, 15) = 5.74, *p* < .05
	27	Voiced, Mismatch vs. Control; language * condition	*F1* (1, 69) = 19.70, *p* < .001	*F2* (1, 15) = 18.21, *p* < .001
	28	Voiced, Dutch; Mismatch vs. Control	*F1* (1, 34) = 17.09, *p* < .001	*F2* (1, 15) = 10.44, *p* < .01
	29	Voiced, English; Mismatch vs. Control	*F1* (1, 35) = 2.42, *p* > .1	*F2* (1, 15) = 1.13, *p* > .1
	30	Voiced, Identity vs. Mismatch; language * condition	*F1* (1, 70) = 19.68, *p* < .001	*F2* (1, 15) = 14.65, *p* < .01

	31	Voiced, Dutch; Identity vs. Mismatch	$F1\ (1, 35) = 4.18, p < .05$	$F2\ (1, 15) < 1$
	32	Voiced, English; Identity vs. Mismatch	$F1\ (1, 35) = 30.00, p < .001$	$F2\ (1, 15) = 8.32, p < .05$
	33[h]	Voiceless; language * condition	$F1\ (2, 138) < 1$	$F2\ (2, 30) < 1$
	34[h]	Condition	$F1\ (2, 138) = 14.69, p < .001$	$F2\ (2, 30) = 10.23, p < .001$
	35	Voiceless; Identity vs. Control	$F1\ (1, 69) = 30.40, p < .001$	$F2\ (1, 15) = 18.86, p < .001$
	36	Voiceless; Mismatch vs. Control	$F1\ (1, 69) = 1.41, p > .1$	$F2\ (1, 15) = 2.61, p > .1$
	37	Voiceless; Identity vs. Mismatch	$F1\ (1, 70) = 15.05, p < .001$	$F2\ (1, 15) = 8.44, p < .05$
V vs. C RT	38	Language * condition * type	$F1\ (2, 140) = 3.62, p < .05$	$F2\ (2, 124) = 5.33, p < .01$

V+C: vowels and consonants; *V*: vowels; *C*: consonants; *V vs. C*: vowels versus consonants.
RT: reaction time; *Corr*: proportion correct
[a, b, c, d, e, f, g, h]: Results with the same superscript were found with a single analysis.

Table 10. Experiment 3, cross-modal priming: English and Dutch listeners' percentage of correct responses and RTs of correct responses for target words in Control, Identity, and Mismatch condition, separately for target words with an /æ/, /ɛ/, voiced (+V) or voiceless final consonant (-V). Examples are given in brackets.

		correct (%)		RT (ms)	
target word	**condition (prime)**	**English**	**Dutch**	**English**	**Dutch**
/æ/ *(lamp)*	**Control *(bike)***	90.6	77.0	613	743
	Identity *(lamp)*	96.4	83.3	579	620
	Mismatch *(lemp)*	93.8	77.1	595	638
/ɛ/ *(desk)*	**Control *(sun)***	97.9	91.7	567	665
	Identity *(desk)*	99.0	90.6	536	646
	Mismatch *(dask)*	99.5	90.6	547	613
+V *(globe)*	**Control *(trade)***	98.4	85.9	588	719
	Identity *(globe)*	99.0	91.1	540	652
	Mismatch *(glope)*	98.4	91.1	608	622
-V *(cheap)*	**Control *(bread)***	97.9	84.3	572	648
	Identity *(cheap)*	99.0	92.2	531	589
	Mismatch *(cheab)*	96.4	93.2	570	620

Consonants

Table 10 also shows the English and Dutch listeners' percentage of correct responses and the RTs of the correct responses in the three conditions, separately for visual targets with voiced and voiceless word-final consonants.

In the analysis of the proportion of correct responses, there was an interaction between native language and condition (Table 9; 14). For the English listeners, there was no main effect of condition (Table 9; 18). Inspection of Table 10 suggests that this may have been due to a ceiling effect. For the Dutch listeners, there was a main effect of condition (Table 9; 17). For the Dutch listeners, the difference between the Identity condition and the Control condition, with a higher proportion correct in the Identity condition than in the Control condition, narrowly missed significance (Table 9; 19). Crucially, for these listeners the proportion correct was higher in the Mismatch condition than in the Control condition (Table 9; 20), and there was no difference between the Identity and the Mismatch conditions (Table 9; 21). Thus, for the Dutch listeners near-words caused facilitation of the corresponding word.

Note that there was no main effect of voicing (Table 9; 15) and there were no interactions involving voicing. Thus, in the proportion of correct responses, there was no asymmetry between items with voiced and voiceless final consonants.

In the analysis of the RTs, there was a three-way interaction among native language, condition, and voicing (Table 9; 22). Therefore, the results for the voiced and voiceless targets were analyzed separately.

For the voiced targets, there was an interaction between native language and condition (Table 9; 24). This interaction was further investigated with ANOVAs comparing the two groups of listeners and two conditions at a time. For the Identity condition versus the Control condition, there was no interaction between native language and condition (Table 9; 25), and RTs were shorter in the Identity condition than in the Control condition (Table 9; 26). For the Mismatch condition versus the Control condition, there was an interaction between native language and condition (Table 9; 27). For the Dutch listeners, RTs were shorter in the Mismatch condition than in the Control condition (Table 9; 28). For the English listeners, there was no difference between RTs in the Mismatch and the Control condition (Table 9; 29). For the Identity condition versus the Mismatch condition, there was an interaction between native language and condition (Table 9; 30). For the Dutch listeners, there was no significant difference between the conditions (Table 9; 31). (Note that for the Dutch listeners RTs were numerically shorter in the Mismatch condition than in the Identity condition.) For the English listeners on the other hand, RTs were shorter in the Identity condition than in the Mismatch condition (Table 9; 32). Thus, for voiced targets, there was facilitation for both Dutch and English listeners in the Identity condition. For the Dutch listeners, there was facilitation in the Mismatch condition, to the same extent as in the Identity condition. For the English listeners, there was no facilitation in the Mismatch condition.

For the voiceless targets, there was no interaction between native language and condition (Table 9; 33). There was a main effect of condition (Table 9; 34), which was further examined in pairwise comparisons. RTs were shorter in the Identity condition than in the Control condition (Table 9; 35). There was no difference between RTs in the Mismatch and the Control condition (Table 9; 36). RTs were shorter in the Identity than in the Mismatch condition (Table 9; 37). Thus, for the voiceless targets, there was significant facilitation in the Identity condition but not in the Mismatch condition, for Dutch and English listeners alike.

Vowels versus consonants

In the analysis of the RTs, there was a significant three-way interaction among native language, condition, and type (vowels vs. consonants) (Table 9; 38). This could be explained by the finding that there was significant facilitation for the Dutch listeners in the Mismatch condition for all the vowel items, but only for a subset of the consonant items.

Phonetic categorization

If listeners based their categorization decision on the nearest lexical representation, this would result in a correct response for real words and an incorrect response for near-words. Therefore, if near-words activated lexical representations more strongly for nonnative listeners than for native listeners, this would result in an interaction between native language and condition (words vs. near-words), with a relatively large number of incorrect responses to near-words for the Dutch listeners. Figure 6 shows that this was the pattern found in the results. Indeed, there was a significant interaction between native language and condition (Table 11; 1). In Figure 6, the results of the words and near-words based on replacement of a vowel and those based on replacement of a consonant are collapsed. Below, the results are discussed in more detail for words and near-words based on a vowel contrast and those based on a consonant contrast separately.

Vowels

Table 12 shows the percentage of correct responses for the vowel targets, separately for words and near-words and for the two original phonemes.

Contrary to expectation, there was no interaction between native language and condition (Table 11; 2). However, the main effect of condition, with more correct responses for real words than for near-words, was significant for the Dutch listeners (Table 11; 4) but not for the English listeners (Table 11; 5).

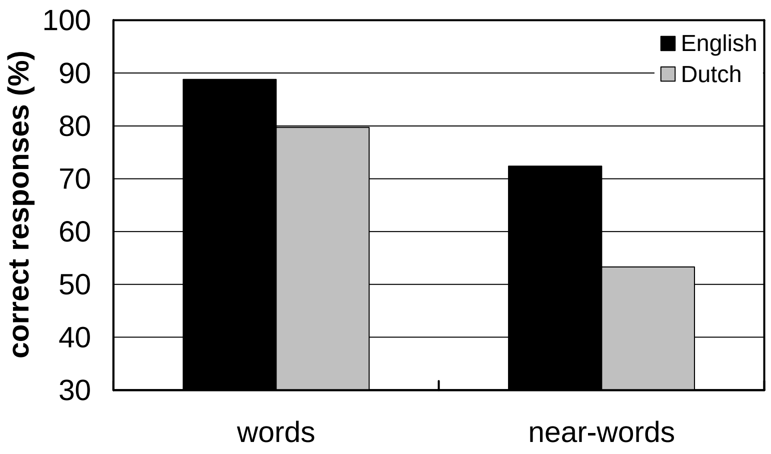

Figure 6. Experiment 3, phonetic categorization: English and Dutch listeners' percentage of correct responses to words and near-words.

Table 11. Experiment 3, phonetic categorization: Analyses.

V+C	1	Language * condition	$F1$ (1, 38) = 6.81, $p < .05$	$F2$ (1, 63) = 8.44, $p < .01$
V	2[a]	Language * condition	$F1$ (1, 38) < 1	$F2$ (1, 30) < 1
	3[a]	Original phoneme * condition	$F1$ (1, 38) = 34.18, $p < .001$	$F2$ (1, 30) = 106.60, $p < .001$
	4	Dutch; condition	$F1$ (1, 19) = 41.71, $p < .001$	$F2$ (1, 30) = 8.15, $p < .01$
	5	English; condition	$F1$ (1, 19) = 33.18, $p < .001$	$F2$ (1, 30) = 3.60, $p = .068$
	6	Dutch; > 50%	t (19) = 8.16, $p < .001$	t (31) = 5.55, $p < .001$
	7	Words; language	$F1$ (1, 38) = 46.89, $p < .001$	$F2$ (1, 30) = 1.96, $p > .1$
C	8	Language * condition	$F1$ (1, 38) = 13.90, $p < .001$	$F2$ (1, 30) = 4.19, $p < .05$
	9	Dutch; condition	$F1$ (1, 19) = 115.55, $p < .001$	$F2$ (1, 30) = 19.35, $p < .001$
	10	English; condition	$F1$ (1, 19) = 80.04, $p < .001$	$F2$ (1, 30) = 6.35, $p < .05$
	11	Dutch; > 50%	t (19) = 8.47, $p < .001$	t (31) = 13.29, $p < .001$
	12	Words; language	$F1$ (1, 38) = 2.94, $p = .095$	$F2$ (1, 30) = 5.20, $p < .05$
V vs. C	13	Language * condition * type	$F1$ (1, 38) = 3.39, $p = .074$	$F2$ (1, 62) = 2.23, $p > .1$

V+C: vowels and consonants; *V*: vowels; *C*: consonants; *V vs. C*: vowels versus consonants.
[a]: Results with the same superscript were found with a single analysis.

Table 12. Experiment 3, phonetic categorization: English and Dutch listeners' percentage of correct responses to words and near-words, separately for items with an /æ/, /ɛ/, voiced (+V) or voiceless final consonant (-V). Examples are given in brackets.

	English	Dutch
word, /æ/ *(lamp)*	80.3	70.7
word, /ɛ/ *(desk)*	96.4	75.6
near-word, /æ/→/ɛ/ *(lemp)*	79.0	54.6
near-word, /ɛ/→/æ/ *(dask)*	53.3	34.8
word, +V *(globe)*	84.6	83.5
word, -V *(cheap)*	94.0	89.1
near-word, +V→-V *(glope)*	80.5	62.3
near-word, -V→+V *(cheab)*	76.9	61.5

Consonants

Table 12 also shows the percentage of correct responses for the consonant targets, separately for words and near-words and for voiced and voiceless original phonemes.

As predicted, in the analysis of the proportion of correct responses, there was a significant interaction between native language and condition (Table 11; 8). The proportion correct was significantly higher for real words than for near-words, both for the Dutch listeners (Table 11; 9) and for the English listeners (Table 11; 10). The interaction between native language and condition showed that the effect of lexicality was stronger for the Dutch listeners than for the English listeners. Thus, near-words activated the nearest lexical representation more strongly for nonnative listeners than for native listeners.

There were no interactions involving original voicing, which means that there was no bias towards either a voiced or a voiceless interpretation of the consonants.

Vowels versus consonants

There was no three-way interaction among native language, condition, and type (Table 11; 13).

Discussion

The results of the cross-modal priming task for the vowel items provided further evidence that near-words cause more activation of lexical representations for nonnative listeners than for native listeners. The results were similar to those of the cross-modal priming task of Experiment 2. For both groups of listeners, recognition of the visual target word was facilitated when it was preceded by auditory presentation of the same word. However, for the Dutch listeners, recognition of the target word was also facilitated when it was preceded by the corresponding near-word, but for the English listeners it was not.

In the phonetic categorization task, for the vowel items, no clear evidence was found that near-words caused more activation of the nearest word form for nonnative listeners than for native listeners. For the Dutch listeners, the proportion of correct responses was lower for near-words than for real words, which showed that lexical representations were active after presentation of near-words. For the English listeners, a similar pattern did not reach significance. However, there was no interaction between native language and condition. In all the previous experiments, evidence was found that vowel near-words caused more activation for nonnative listeners than for native listeners. The difference with the phonetic categorization tasks in Experiments 1 and 2 seems to be that the lexical effect for the English listeners was larger in the present experiment. This may be due to the way the stimuli of this experiment were constructed, with a word boundary within each item, which may have made the vowels perceptually ambiguous for the English listeners. Target phonemes often followed the word boundary directly, or else they were the second (and in one case the third) phoneme after the word boundary. The word boundary within the stimulus may have hindered the recognition of the vowel more for the English listeners than for the Dutch listeners, as the Dutch listeners may not have been sensitive to particular acoustic cues anyway. Indeed, whereas the English listeners categorized the vowels in real words significantly more accurately than the Dutch listeners in Experiments 1 (Table 3; 8) and 2 (Table 7; 7), this difference was not significant in the present experiment (Table 11; 7). In the cross-modal priming task, using the same stimuli, evidence was found that presentation of a near-word did lead to more lexical activation for the Dutch listeners than for the English listeners.

For the consonant items, the results of both tasks showed that near-words caused more activation of lexical representations for nonnative listeners than for native listeners. In the cross-modal priming task, for the English listeners, recognition of the visual target word was facilitated when it was preceded by auditory presentation of the same word, but not after presentation of a near-word. For the Dutch listeners, recognition of the visual target was facilitated after presentation of the real word but also after presentation of the corresponding near-word. Facilitation after presentation of a near-word was found in the proportion of correct responses for all consonant items, and in the RTs for the target words with a voiced final consonant.

In the phonetic categorization task, the effect of lexical status, with a lower proportion of correct responses to near-words than to real words, was larger for the Dutch listeners than for the English listeners. Thus, categorization responses to a near-word were more often consistent with the nearest word form for nonnative listeners than for native listeners. This indicates that there was more lexical activation upon presentation of near-words for nonnative listeners than for native listeners. The consonant targets always occurred at the end of the stimulus. As these target phonemes did not occur so closely after the word boundary as the vowel targets, preceding acoustic cues signaling the final consonant's identity may have been unaffected.

In the cross-modal priming task, near-words with an /æ/ and those with an /ɛ/ caused a similar amount of lexical activation. However, for words with voiced and voiceless final consonants, there was an asymmetry for the Dutch listeners but not for the English listeners. For the Dutch listeners, RTs to target words with voiced consonants showed that auditory presentation of words and near-words facilitated the recognition of the visual target word to the same extent. Thus, presentation of a near-word led to as much lexical activation as presentation of a real word, and the discrepancy between the near-word and the word did not lead to deactivation of the lexical representation. For the target words with voiceless consonants, recognition was significantly faster than in the Control condition after presentation of a real word, but not after presentation of a near-word. However, although there was no significant facilitation after presentation of these near-words in the RTs, the analysis of the proportion correct provided evidence that some activation of the voiceless targets did remain after presentation of the near-words: the proportion of correct responses to visual target words was higher after presentation of near-words than after presentation of unrelated words. The effect of near-words on the proportion of correct responses was as large as the effect of real words, and occurred for voiced and voiceless items alike. Thus, the discrepancy between a near-word with a voiced final consonant and a word with a voiceless final consonant led to some deactivation of the lexical representation (as the RTs showed), but some activation of the lexical representation remained (as the proportion correct showed).

Thus, for the Dutch listeners, there was an asymmetry between the near-word *glope* activating *globe*, and the near-word *cheab* activating *cheap*. The word *globe* and the near-word *glope* activated *globe* to the same extent. On the other hand, the near-word *cheab* did activate *cheap*, but not as much as the word *cheap* itself did. For the English listeners, the lexical representations were not active after either near-word.

General discussion

There is more spurious lexical activation in the comprehension of nonnative speech than in the comprehension of native speech. The experiments presented here provide very clear and consistent results. They provide converging evidence that near-words lead to more lexical activation for nonnative listeners than for native listeners.

The amount of lexical activation after presentation of a near-word was very high for the nonnative listeners. For the Dutch listeners, in many cases, presentation of a near-word caused as much activation of the nearest word form as presentation of the word itself did. This was investigated with two cross-modal priming experiments, which provide an on-line measure of the amount of lexical activation upon presentation of words and near-words.

Experiments 2 and 3 both showed that presentation of near-words based on a vowel replacement activated the nearest word form as much as presentation of the real word did. Words and near-words like *deaf* and *daf* activated the lexical representation of *deaf* to the same extent. In Experiment 3, a similar result was found for near-words based on the replacement of a voiced final consonant with a voiceless one. Words and near-words like *globe* and *glope* activated the lexical representation of *globe* to the same extent. Thus, the discrepancy between a near-word and a word did not lead to any deactivation of the lexical representation. Presentation of the crucial phoneme did not lead to bottom-up inhibition but, on the contrary, to a further activation of the lexical representation. The amount of spurious lexical activation caused by a near-word determines how much it complicates the comprehension process. If spurious lexical competitors were only weakly activated, they would be relatively easy to inhibit and their influence on the comprehension process would be limited. However, these results show that the level of spurious lexical activation caused by near-words is very high, namely often as high as the level of lexical activation caused by real words. Therefore, spurious lexical activation may not be easy to overcome, and this may greatly complicate the comprehension of nonnative speech.

One likely cause for spurious lexical activation in the comprehension of nonnative speech is perceptual ambiguity. For example, if listeners cannot accurately distinguish between *lamp* and *lemp*, presentation of *lemp* may lead to the activation of the lexical representation of *lamp*. It is argued that perceptual ambiguity played a role for the Dutch listeners for the items based on the /æ/-/ɛ/ contrast, but not for the items based on the final obstruent voicing contrasts. Nevertheless, there was increased lexical activation for the Dutch listeners compared to the native listeners after near-words based on the final obstruent voicing contrasts. Thus, spurious lexical activation is not restricted to perceptually ambiguous stimuli.

The word-final consonant voicing contrast is not perceptually ambiguous for Dutch listeners. Nevertheless, Experiments 1 and 3 showed that near-words based on the word-final voicing contrast caused more spurious lexical activation for Dutch listeners than for English listeners. In Experiment 3, spurious lexical activation occurred in an asymmetrical way. Although near-words of both types caused more activation of the nearest word form for nonnative listeners than for native listeners, for the Dutch listeners, near-words like *glope* caused more activation of words like *globe* than near-words like *cheab* did of words like *cheap*. As proposed in the Introduction, one possible explanation could be that the Dutch listeners found it easier to disregard the voicing of a final voiceless obstruent for lexical access than that of a voiced obstruent. As voiced obstruents are not allowed at the end of words in isolation in Dutch, they may be marked and difficult to ignore for Dutch listeners. Another explanation could be that the Dutch listeners' lexical representations differed from native listeners' representations. Dutch speakers may mispronounce *globe* as *glope* more often than they mispronounce *cheap* as *cheab*, and this might be reflected in the mental lexicon of Dutch listeners, who are often exposed to these mispronunciations. Note that the differential effect of *glope* and *cheab* cannot be accounted for by an asymmetry in the perception of voiced and voiceless phonemes. If such an asymmetry existed, it should have surfaced even more clearly in a phonetic categorization task. In the phonetic categorization tasks of Experiments 1 and 3, no bias towards a voiced (or voiceless) interpretation of word-final obstruents was found. Similarly, Broersma (2005, submitted) and Van Wieringen (1995) found no bias in Dutch listeners' phonetic categorization of voiced and voiceless final obstruents.

Although the Dutch listeners' lexical representations may sometimes have been nonnative-like, lexical representations were not always inaccurate with respect to the contrasts under study. All experiments showed that the Dutch listeners knew which phoneme a particular word contained with a level of accuracy above chance. In the auditory lexical decision task of Experiment 1, both Dutch and English listeners gave more "yes" responses to real words than to near-words. They did not treat words and near-words as equally acceptable forms of the same word. In the phonetic categorization tasks of Experiments 1, 2, and 3, an effect of lexicality was found for the Dutch listeners. For these listeners, the proportion of correct responses was lower for near-words than for words. As the Dutch listeners were inclined to interpret the target phoneme such that it was consistent with the standard pronunciation of the word, they must have known which phoneme the word contained. Thus, the Dutch listeners' lexical representations must have been correctly coded for the /æ/-/ɛ/ contrast and the consonant voicing contrasts in a majority of the cases.

Although the central question examined in this study concerns the occurrence of spurious lexical activation, some other observations are also worth mentioning.

First, the results from the three phonetic categorization experiments showed that in the categorization of the /æ/-/ɛ/ contrast, both the Dutch and the English listeners had a bias towards "/ɛ/" responses. In the phonetic categorization tasks of Experiments 1, 2, and 3, there was an asymmetry between the proportion of correct responses to the two vowels, which appeared in all these experiments from an interaction between condition and original phoneme (Table 3; 4, Table 7; 3, Table 11; 3). The proportion of correct responses to items containing /ɛ/ was higher than that to items containing /æ/. The proportion correct to words like *desk* was higher than to words like *lamp*, and the proportion correct to near-words like *lemp* was higher than to near-words like *dask*. This bias towards /ɛ/ was consistently and significantly present in the three phonetic categorization tasks and for Dutch and English listeners alike. Broersma (submitted) found a similar bias in Dutch and English listeners' phonetic categorization of the /æ/-/ɛ/ contrast. The explanation may be that the vowel /ɛ/ occurs more frequently in English than the vowel /æ/. In the CELEX lexical database (Baayen et al., 1995), the number of word forms containing an /ɛ/ is similar to those containing an /æ/ (14,006 vs. 14,230), but the mean logarithmic lemma frequency per million of words containing an /ɛ/ is higher than the mean frequency of words containing an /æ/ (0.34 vs. 0.26).

The bias towards perception of /ɛ/ had an impact on the recognition of words containing /æ/ or /ɛ/. In the lexical decision task of Experiment 1 and the cross-modal priming tasks of Experiments 2 and 3, upon presentation of a word or a near-word, the participants found it easier to recognize words with an /ɛ/ than words with an /æ/. In the auditory lexical decision task of Experiment 1, there was an asymmetry between the words and near-words with an /ɛ/ and those with an /æ/ in the original word, showing as a main effect of original phoneme (Table 1; 5). Words with an /ɛ/ and their corresponding near-words were judged as a word more often than words with an /æ/ and their corresponding near-words. Thus, there were more "yes" responses to items like *desk* and *dask* than to items like *lamp* and *lemp*. This can be explained by the bias found in the phonetic categorization experiments. If both *desk* and *dask* were relatively often perceived as *desk*, whereas both *lamp* and *lemp* were relatively often perceived as *lemp*, this would lead to a high percentage of "yes" responses to *desk* and *dask*, and a low percentage of "yes" responses to *lamp* and *lemp* indeed. In the cross-modal priming task of Experiment 3, the overall proportion of correct responses was higher (Table 9; 12) and the RTs of correct responses were shorter (Table 9; 3) for visual target words with an /ɛ/ than for those with an /æ/. Similar but non-significant effects were found in the cross-modal priming task of Experiment 2, for the proportion correct (Table 5; 12) and for the RTs (Table 5; 11). This can also be explained by the bias found in the phonetic categorization experiments. If the Identity and Mismatch primes for *desk*, namely *desk* and *dask*, were relatively often perceived as *desk*, whereas the primes for *lamp*, namely *lamp* and *lemp*, were relatively often perceived as *lemp*, there should indeed be more facilitation for the recognition of the words containing an /ɛ/ than of those containing an /æ/. Note also that this effect of the bias towards perception of /ɛ/ is further evidence that the Dutch listeners' lexical representations were correctly specified for the vowels /æ/ and /ɛ/ in a majority of the cases.

Second, some evidence was found that there was more lexical activation after a vowel mismatch than after a consonant mismatch. This pattern emerged from all tasks in which both near-words based on a vowel replacement and near-words based on a consonant replacement were used, for Dutch and English listeners alike. In the auditory lexical decision task of Experiment 1, there was an interaction between condition (words vs. near-words) and type (vowel items vs. consonant items) (Table 1; 17). For real words, the proportion of "yes" responses was lower for vowel items than for consonant items (Table 1; 18). However, for the near-words, the proportion of "yes" responses was higher for vowel items than for consonant items (Table 1; 19). Thus, near-words based on the replacement of a vowel were more often accepted as a word than near-words based on the replacement of a consonant. In the phonetic categorization task of Experiment 1, there was also an interaction between condition and type (Table 3; 16). Both for words (Table 3; 17) and near-words (Table 3; 18), the proportion of correct responses was higher for consonant items than for vowel items, but the difference was larger for the near-words than for the words. Thus, near-words based on a vowel replacement caused more lexical activation than near-words based on a consonant replacement. The results of the phonetic categorization task of Experiment 3 showed a similar but non-significant pattern. In the cross-modal priming task of Experiment 3, there was no significant interaction between condition and type either. However, numerically, near-words based on a vowel replacement caused more facilitation in the RTs than near-words based on a consonant replacement. Together, the results of these experiments suggest that there was more lexical activation after a vowel near-word than after a consonant near-word, both for native and nonnative listeners. These results are in line with the idea that vowels may constrain lexical selection less tightly than consonants do (Cutler, Sebastián-Gallés, Soler-Vilageliu, & Van Ooijen, 2000).

Although there was more spurious lexical activation after a vowel near-word than after a consonant near-word both for native and for nonnative listeners, the difference between vowel and consonant near-words was sometimes larger for the Dutch listeners than for the English listeners. This was shown by significant three-way interactions between native language, condition, and type in the phonetic categorization task of Experiment 1 (Table 3; 15) and in the cross-modal priming task of Experiment 3 (Table 9; 38). This could suggest that perceptual ambiguity (which was only expected to play a role for the vowels and for the nonnative listeners) might be the most effective cause for spurious lexical activation. However, no three-way interactions were found in the lexical decision task of Experiment 1 (Table 1; 16) and in the phonetic categorization task of Experiment 3 (Table 11; 13).

Third, the results from the phonetic categorization tasks provide some insight in the Dutch listeners' categorization of the /æ/-/ɛ/ and the word-final consonant voicing contrasts. As lexical effects played an important role in the phonetic categorization tasks, no strong conclusions can be drawn about perceptual accuracy. However, as lexical effects enhanced the categorization of real words and hindered the categorization of near-words, together the data do provide some information. When the responses to words and near-words were collapsed, the Dutch listeners' categorization of the vowels reached a level of accuracy which was significantly above chance level (50 % correct) in all experiments (Experiment 1: Table 3; 7, Experiment 2: Table 7; 6, Experiment 3: Table 11; 6). This is in line with the findings of Broersma (2005, submitted). Similarly, the Dutch listeners categorized the consonant voicing contrasts with a level of accuracy above chance (Experiment 1: Table 3, 13, Experiment 3: Table 11, 11), in line with Broersma (2005, submitted), who found that Dutch listeners categorized these contrasts as accurately as English listeners did.

Fourth, not surprisingly, there were differences between the accuracy of the English listeners and the Dutch listeners in the (English) experiments described in this paper. In addition to the finding that near-words caused more spurious lexical activation for nonnative listeners than for native listeners, there were also differences in the way both groups responded to real word stimuli. The following differences were found between the English and the Dutch listeners. In the auditory lexical decision task of Experiment 1, the proportion of correct responses to words was higher for English listeners than for Dutch listeners (Table 1; 3). In the cross-modal priming tasks of Experiments 2 (Table 5; 2) and 3 (Table 9; 13 and 16), the RTs of the correct responses were shorter for the English listeners than for the Dutch listeners. In Experiment 3, the proportion of correct responses to visual target words was higher for the English listeners than for the Dutch listeners (Table 9; 4 and 23). A similar effect missed significance in Experiment 2 (Table 5; 13). In the phonetic categorization tasks of Experiments 1 (Table 3; 8) and 2 (Table 7; 7), for the vowel targets in a real word, the proportion of correct responses of the English listeners was higher than that of the Dutch listeners. This is in line with the finding of Broersma (2005, submitted) that Dutch listeners categorized the /æ/-/ɛ/ contrast less accurately than English listeners did. However, in the phonetic categorization tasks of Experiments 1 (Table 3; 14) and 3 (Table 11; 12), for the consonant targets in real words, the Dutch listeners' proportion of correct responses did not differ from that of the English listeners. This is also in line with the finding of Broersma (2005, submitted) that Dutch listeners categorized final voiced and voiceless consonants as accurately as English listeners did.

In summary, this study provides converging evidence, obtained with three different experimental paradigms, that near-words cause more spurious lexical activation for nonnative listeners than for native listeners. For nonnative listeners, the level of lexical activation after presentation of a near-word was very high, and often as high as the level of activation after presentation of a real word. Spurious lexical activation may result from perceptual ambiguity of nonnative phonemes. However, this study provides evidence that spurious lexical activation may also occur when phoneme recognition is uncompromised. In this study, Dutch listeners with a high level of proficiency in English were tested. The occurrence of spurious lexical activation may be even more common for listeners with a lower level of proficiency in their nonnative language.

The failure to deactivate spurious lexical competitors is likely to make speech comprehension more difficult for nonnative listeners. According to the TRACE (McClelland & Elman, 1986) and Shortlist (Norris, 1994) models of speech comprehension, all activated word forms compete for recognition. Although lexical competition is a necessary part of speech comprehension, it also makes the recognition of the intended words more difficult. Lexical competitors diminish the activation of the intended word through lateral inhibition. Target words are more difficult to recognize in a speech input that is also consistent with a lexical competitor (McQueen et al., 1994). The more competitors are active, the harder it is to recognize the speech input correctly (Luce et al., 1990; Norris et al., 1995; Vitevitch & Luce, 1999; Vroomen & De Gelder, 1995).

The results from this study show that spurious lexical competitors which are not active in the lexicon of native listeners are active for nonnative listeners. Thus, when native and nonnative listeners listen to the same speech input, more lexical competitors may be active for the nonnative listeners than for the native listeners. Like all lexical competitors, spurious lexical competitors will compete for recognition and hinder the selection of the intended words. Therefore, spurious lexical activation may hinder the recognition of nonnative speech considerably. This may be one of the reasons why speech comprehension is so much more difficult in a nonnative language than in the native language.

References

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

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.

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

Broersma, M. (2005). Perception of familiar contrasts in unfamiliar positions. *Journal of the Acoustical Society of America, 117*, 3890-3901. *Chapter 2 of this dissertation.*

Broersma, M. (submitted). Competition increase in nonnative listening. *Language and Speech. Chapter 4 of this dissertation*.

Connine, C. M., Blasko, D. G., & Wang, J. (1994). Vertical similarity in spoken word recognition: Multiple lexical activation, individual differences, and the role of sentence context. *Perception and Psychophysics, 56*, 624-636.

Cutler, A. (2005). *The lexical statistics of word recognition problems caused by L2 phonetic confusion*. Proceedings of the 9th European Conference on Speech Communication and Technology, Lisbon, September 2005.

Cutler, A., & Otake, T. (2004). *Pseudo-homophony in non-native listening*. Poster presented at the 75th meeting of the Acoustical Society of America, New York.

Cutler, A., Sebastián-Gallés, N., Soler-Vilageliu, O., & Van Ooijen, B. (2000). Constraints of vowels and consonants on lexical selection: Cross-linguistic comparisons. *Memory and Cognition, 28*, 746-755.

Cutler, A., Weber, A., & Otake, T. (in press). Asymmetric mapping from phonetic to lexical representations in second-language listening. *Journal of Phonetics*.

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.

Eimas, P. D., Marcovitz Hornstein, S. B., & Payton, P. (1990). Attention and the role of dual codes in phoneme monitoring. *Journal of Memory and Language, 29*, 160-180.

Frauenfelder, U. H., Scholten, M., & Content, A. (2001). Bottom-up inhibition in lexical selection: Phonological mismatch effects in spoken word recognition. *Language and Cognitive Processes, 16*, 583-607.

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.

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

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.

Marslen-Wilson, W. D. (1987). Functional parallelism in spoken word-recognition. *Cognition, 25*, 71-102.

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

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

McQueen, J. (2004). Speech perception. In K. Lamberts & R. Goldstone (Eds.), *The Handbook of Cognition* (pp. 255-275). London: Sage Publications.

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.

McQueen, J. M., Cutler, A., Briscoe, T., & Norris, D. (1995). Models of continuous speech recognition and the contents of the vocabulary. *Language and Cognitive Processes, 10*, 309-331.

Norris, D. (1994). Shortlist: A connectionist model of continuous speech recognition. *Cognition, 52*, 189-234.

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. (2000). Merging information in speech recognition: Feedback is never necessary. *Behavioral and Brain Sciences, 23*, 299-325.

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.

Pitt, M. A., & Samuel, A. G. (1993). An empirical and meta-analytic evaluation of the phoneme identification task. *Journal of Experimental Psychology: Human Perception and Performance, 19*, 699-725.

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.

Shillcock, R. (1990). Lexical hypotheses in continuous speech. In G. T. M. Altmann (Ed.), *Cognitive Models of Speech Processing: Psycholinguistic and Computational Perspectives* (pp. 24-49). Cambridge, MA: MIT.

Strange, W. (Ed.). (1995). Speech Perception and Linguistic Experience: Issues in Cross-Language Research. Baltimore: York Press.

Tabossi, P., Burani, C., & Scott, D. (1995). Word identification in fluent speech. *Journal of Memory and Language, 34*, 440-467.

Van Wieringen, A. (1995). *Perceiving Dynamic Speechlike Sounds: Psycho-acoustics and Speech Perception.* Unpublished doctoral dissertation, Universiteit van Amsterdam.

Vitevitch, M. S., & Luce, P. A. (1999). Probabilistic phonotactics and neighborhood activation in spoken word recognition. *Journal of Memory and Language, 40*, 374-408.

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.

Vroomen, J., & De Gelder, B. (1997). Activation of embedded words in spoken word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 23*, 710-720.

Weber, A., & Cutler, A. (2004). Lexical competition in non-native spoken-word recognition. *Journal of Memory and Language, 50*, 1-25.

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

Appendix A

Experimental stimuli used in Experiments 1 and 3.

Words from Experiment 1 served as Identity primes in the cross-modal priming task of Experiment 3, near-words as Mismatch primes.

In the cross-modal priming task of Experiment 3, Identity primes also served as visual targets. Fictitious geographic names in the carrier fragments are indicated with a *.

Primes:			Carrier fragments:		
Identity	Mismatch	Control	Identity	Mismatch	Control
/æ/					
bank	benk	skeep	superb ankle	superb enclave	glass keep
black	bleck	tring	superb lack	superb lek	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 ratch	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	splesh	tring	nice Plashterville*	nice Pleshterville*	hot ring
thank	thenk	bike	fourth ankle	fourth enclave	prefab icon
trap	trep	skeep	that rap	that rep	glass keep
/ɛ/					
bench	banch	sun	superb enchilada	superb anchovy	this undertone
breast	brast	trick	prefab restaurant	prefab raster	that rick
breath	brath	skeep	superb Reathly*	superb Rathly*	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

Primes:			Carrier fragments:		
Identity	Mismatch	Control	Identity	Mismatch	Control
desk	dask	sun	bad escort	bad asking	this undertone
dress	drass	grain	good rescue	good raster	big rain
fresh	frash	buse	half Reshdale*	half Rashdale*	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	swat	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
/z/					
cheese	cheece	friend	much easier	much Easter fun	brief rendering
news	newce	friend	does Lynn use	in use	brief rendering
phrase	phrace	crup	safe raise	safe race	weak rupture
praise	praice	friend	sharp raise	sharp race	brief rendering
/s/					
choice	choise	crup	reach Oice Lake*	reach Oise Lake*	weak rupture
kiss	kiz	cup	weak ischial joint	weak Israel	weak upside
nurse	nurze	cup	in Erse	Lynn errs	weak upside
voice	voise	crup	leave Oice Lake*	leave Oise Lake*	weak rupture
/v/					
dive	dife	spend	good ivy	good eyeful	Swiss pendant
groove	groof	spend	big rooves	big roofs	Swiss pendant
move	moof	fres	calm Oovington*	calm Oofington*	safe rescue
shave	shafe	fres	posh aviator	posh aphid	safe rescue
/f/					
dwarf	dwarve	prup	bad warfare	bad war victim	sharp rupture
laugh	lauve	prup	until after	until Arvin	sharp rupture
scarf	scarve	prup	ox calf	does the ox calve	sharp rupture
stiff	stiv	spend	least iffy	east Ivrea	Swiss pendant

Primes:			Carrier fragments:		
Identity	Mismatch	Control	Identity	Mismatch	Control
/b/					
cube	cupe	lart	dark U-boat	dark upas tree	evil artist
globe	glope	trade	big lobe	big lope	wet radar
rub	rup	lart	clear ubble-gubble	clear uplift	evil artist
tube	tupe	lart	wet U-boat	wet upas tree	evil artist
/p/					
cheap	cheab	bread	each epoch	each ebola outbreak	superb readiness
sharp	sharb	bread	lush arpeggio	lush arboreal	superb readiness
sheep	sheeb	fres	lush epoch	lush ebola outbreak	safe rescue
ship	shib	bread	finish Ipswich	finish ibidem	superb readiness
/d/					
beard	beart	flime	arab eardrum	arab ear trumpet	stiff lime
blade	blate	flime	the grebe laid	is the grebe late	stiff lime
glide	glite	flime	the pig lied	big light	stiff lime
proud	prout	cup	the chap rowed	trap route	weak upside
/t/					
flight	flied	care	half light	half lied	sick air
skirt	skird	care	this is curt	this is curd	sick air
smart	smard	care	Swiss martyr	Swiss Mardi Gras	sick air
spit	spid	pide	base pity	base piddle	cheap idol

Appendix B

Experimental stimuli used in Experiment 2.

Identity primes also served as visual targets.

Primes:			Carrier fragments:		
Identity	Mismatch	Control	Identity	Mismatch	Control
/æ/					
cat	ket	pock	catalogue	kettle	pocket
damn	dem	al	damage	democrat	albatross
lamb	lem	bal	laminate	lemon	balance
pant	pent	synth	pantomime	pentagram	synthesize
tack	teck	fing	tactic	textile	finger
tan	ten	bul	tantrum	tentacle	bulletin
/ɛ/					
chess	chas	an	chestnut	chastity	anarchism
deaf	daf	hov	definite	daffodil	hovercraft
dress	dras	stupe	dressage	drastic	stupid
edge	adge	whis	educate	agile	whistle
egg	ag	lune	egotist	agony	lunatic
shed	shad	vir	schedule	shadow	virgin

Competition increase in nonnative listening

Chapter 4

Broersma, M. (submitted). Competition increase in nonnative listening. Language and Speech.

Abstract

Two cross-modal priming experiments investigated the increased activation of lexical competitors in nonnative as opposed to native listening. Experiment 1 investigated Dutch and English listeners' perception of English words with partially overlapping onsets (e.g., *daffodil* - *deficit*). Onsets differed in the /æ/-/ɛ/ contrast, which was perceptually ambiguous for the nonnative listeners. Presentation of the onset of one word activated the paired word more for nonnative listeners than for native listeners. Experiment 2 investigated perception of minimal pairs, differing either in the /æ/-/ɛ/ contrast (e.g., *flash* - *flesh*) or in a word-final consonant voicing contrast which was also nonnative but easy to distinguish for the nonnative listeners (e.g., *robe* - *rope*). Presentation of one word facilitated recognition of the same word for native and nonnative listeners alike. It inhibited recognition of the minimally different word for the native listeners, but produced neither facilitation nor inhibition for the nonnative listeners. This pattern occurred for both the vowel and the voicing contrast. The results suggest that partially overlapping words can remain active longer for nonnative listeners, causing an increase of lexical competition in nonnative compared with native listening.

Introduction

The comprehension of speech in a second language is difficult in many ways. Some of the difficulties involved, like the listener's unfamiliarity with a word, the confusability of two speech sounds, or the inability to segment the incoming speech stream into separate words, may be clearly noticeable to the listener. Other difficulties may be less noticeable, but they may nevertheless severely complicate the speech perception process. One of these unnoticed but serious difficulties in nonnative listening is the increased activation of lexical competitors.

The activation of multiple lexical representations is a necessary part of speech comprehension, both in the native language and in second languages. The number of phonemes that a language has at its disposal is limited, whereas the number of possible words built up of those phonemes is very large. One consequence is that a large majority of polysyllabic words have shorter words embedded in them (McQueen, Cutler, Briscoe, & Norris, 1995), another that many words partially overlap with others. The degree of overlap varies from word pairs sharing one phoneme, to minimal pairs, overlapping in all phonemes but one. When listeners hear a word containing an embedded word, both the longer word and the embedded word are activated in the mental lexicon (Davis, Marslen-Wilson, & Gaskell, 2002; Salverda, Dahan, & McQueen, 2003). When listeners hear a word which partially overlaps with another word, the partially overlapping lexical competitor is also activated (Zwitserlood, 1989). Current models of speech recognition agree on this notion of multiple lexical activation (for a review see McQueen, 2004).

Activated word forms actively compete for recognition (McQueen, Norris, & Cutler, 1994). As lexical competition may lead to the deactivation of competitors and to the selection of the target word, it is conducive to speech comprehension. However, the other side of the coin is that words are harder to recognize when more lexical competitors are active. Thus, it is more difficult to recognize words when the number of words that partially match the input is larger (Norris, McQueen, & Cutler, 1995; Vroomen & De Gelder, 1995) and when they have more lexical neighbors, and thus more lexical competitors (Luce, Pisoni, & Goldinger, 1990).

This threatens to be a severe problem in nonnative listening. There is a growing body of evidence that there is more activation of lexical competitors in nonnative listening than in native listening. Increased lexical activation in nonnative listening has been found to occur in different ways. First, non-words may be perceived as real words. Second, words which partially overlap with the input may remain active after a segmental mismatch. Third, minimal pairs may be perceived as homophones.

Evidence that non-words are sometimes perceived as real words comes from Broersma (2002, submitted), who studied the perception of British English by Dutch listeners with a high level of proficiency in English as a second language. Listeners were presented with words and non-words which only differed from these words in the vowels /æ/ and /ɛ/, which are difficult to distinguish for Dutch listeners (e.g., *lamp - lemp*). Non-words were excised from a single carrier word or a two-word context (e.g., *eviL EMPire*), as that is how they may occur in normal speech. Auditory lexical decision, cross-modal priming, and phonetic categorization experiments showed that presentation of a non-word caused more activation of the word form for nonnative listeners than for native English listeners.

Further evidence about the perception of non-words comes from Sebastián-Gallés, Echeverría, and Bosch (2005). They studied the perception of Catalan by highly fluent Spanish-Catalan early bilinguals whose native language was either Spanish or Catalan, and by bilinguals who acquired both languages simultaneously. Catalan words were used that contained either the vowel /e/ or /ɛ/, which are difficult to distinguish for Spanish-dominant listeners. On the basis of these words, non-words were created by replacing /e/ by /ɛ/ and vice versa. Words and non-words were presented for auditory lexical decision. The results showed that the early Spanish-dominant bilinguals treated the non-words more often like words than the Catalan-dominant bilinguals did. Even the simultaneous bilinguals responded to non-words as if they were words more often than the Catalan-dominant bilinguals did.

Weber and Cutler (2004) showed that partially overlapping competitors remain active longer in nonnative listening than in native listening. They studied the perception of British English by Dutch listeners from the same population as Broersma (2002, submitted). The study made use of word pairs with onsets which overlapped except for either the /æ/-/ɛ/ or the /aɪ/-/eɪ/ contrast, which are perceptually ambiguous for Dutch listeners. In an eye-tracking experiment, it was found that presentation of a word like *panda* led to the activation of its partially overlapping competitor *pencil* for the Dutch listeners but not for native listeners of English.

Similar results were found in an eye-tracking experiment for Japanese listeners' perception of their second language English (Cutler, Weber, & Otake, in press). Word pairs had overlapping onsets except for the difficult to distinguish /r/-/l/ contrast. The word *rocket* was found to activate *locker*. Thus, activation of the competitor remained after the /r/-/l/ mismatch.

Finally, there is evidence that nonnative listeners sometimes treat minimal pairs as homophones. Pallier, Colomé, and Sebastián-Gallés (2001) studied the perception of Catalan by early Spanish-Catalan bilinguals whose native language was either Spanish or Catalan (similar to those tested by Sebastián-Gallés et al., 2005). Minimal pairs of Catalan words were used which differed in the contrast /e/-/ɛ/, /o/-/ɔ/, or /s/-/z/, which are difficult to perceive for Spanish-dominant listeners (e.g., */nɛtə/ - /netə/*). In an auditory repetition-priming task, facilitation of the recognition of a word due to the previous presentation of the paired word was investigated. Whereas the Catalan-dominant listeners did not show any facilitation for presentations of minimal pairs, the Spanish-dominant listeners did. For the Spanish-dominant listeners, facilitation upon presentation of the second item of a minimal pair was similar in size to that after repetition of the same word. Thus, they treated the minimal pairs as homophones.

Similarly, Cutler and Otake (2004) present evidence from auditory repetition priming that English minimal pairs differing in the /r/-/l/ contrast (e.g., *write – light*) activate one another for Japanese listeners, and minimal pairs differing in the /æ/-/ɛ/ contrast (e.g., *cattle – kettle*) activate one another for Dutch listeners.

Thus, there is evidence that there is more competitor activation in nonnative than in native speech perception. Non-words were perceived as words, partially overlapping words remained ambiguous longer, and minimal pairs were perceived as homophones. This may be the result of the perceptual ambiguity of nonnative speech sounds. A large body of evidence (e.g., Strange, 1995) shows that the perception of particular phoneme contrasts can be difficult for nonnative listeners. Connine, Blasko, and Wang (1994) showed that perceptually ambiguous stimuli which were equally consistent with two words (e.g., *dent* and *tent*) activated two lexical representations to the same extent. Similarly, perceptually ambiguous phoneme contrasts in a second language may lead to increased competitor activation. Indeed, in the studies described above, non-words, partially overlapping words, and minimal pairs were based on phoneme contrasts which were perceptually ambiguous for the nonnative listeners. However, Broersma (2002, submitted) found that increased competitor activation also occurred for non-words which were not based on a difficult to distinguish contrast. This is further investigated in the present paper.

The present study examines the occurrence of increased lexical activation in nonnative listening with an experimental paradigm particularly suited for the task. The cross-modal priming paradigm (see Zwitserlood, 1996) has been shown to be a powerful tool to provide insight into lexical activation and competition. It has been used successfully to investigate effects of perceptual ambiguity (Connine et al., 1994) and of mismatch (Soto-Faraco, Sebastián-Gallés, & Cutler, 2001) on the amount of lexical activation. Further, it provides insight into the strength of lexical competition (Gaskell & Marslen-Wilson, 2002; Vroomen & De Gelder, 1995). It has been used successfully to investigate the processing of homophones in the perception of native language speech (Onifer & Swinney, 1981). Finally, the cross-modal priming paradigm has proven to be a particularly good task to study the perception of non-words in native and nonnative listening (Broersma, 2002, submitted). Therefore, the present study made use of the cross-modal priming task to assess lexical activation and competition in the perception of nonnative speech. It investigated the processing of word pairs with partially overlapping onsets (Experiment 1) and of minimal pairs (Experiment 2).

Experiment 1

Experiment 1 was designed to investigate Dutch and English listeners' perception of English words with partially overlapping onsets. Word pairs had a similar onset, except for the vowels /æ/ and /ɛ/ (e.g., *daffodil – deficit*). The experiment consisted of two tasks, a cross-modal fragment priming task and a phonetic categorization task. The phonetic categorization task was included to obtain information about the nonnative listeners' ability to distinguish the /æ/-/ɛ/ contrast.

The /æ/-/ɛ/ contrast was expected to be perceptually difficult for the Dutch listeners. Dutch has only one phoneme in the phonetic space of English /æ/ and /ɛ/. Therefore, according to the Perceptual Assimilation Model (Best, 1994), Dutch listeners are expected to initially assimilate both phonemes to a single native category, which would make the perception of the contrast difficult. Dutch listeners were found to have difficulty distinguishing between these vowels. In a phonetic categorization experiment in which the vowels were presented in a non-word context, Dutch listeners categorized the vowels with a level of accuracy which was amply above chance but significantly lower that that of native listeners of English (Broersma, 2005).

Method

Participants

Seventy-two native speakers of Dutch and 72 native speakers of British English took part. The Dutch participants had a high level of proficiency in English as a second language. They had received on average 7 years of English instruction in primary and secondary education. The English participants did not know any Dutch. 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. None reported any hearing loss, visual loss, or reading disability. All were volunteers and received a small fee for participation.

Materials

For the cross-modal priming task, 24 pairs of trisyllabic English words with stress on the first syllable were selected as visual target words. For each pair, the first parts of the two words, up to and including the vowel of the second syllable, were identical, except that one word had an /æ/ in the first syllable and the other an /ɛ/ (e.g., *daffodil – deficit*). The mean logarithmic lemma frequency per million, determined with the CELEX lexical database (Baayen, Piepenbrock, & Gulikers, 1995), was 0.98 for the words with an /æ/ and 0.89 for those with an /ɛ/. For each pair, a phonologically and semantically unrelated trisyllabic word was selected.

All words 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 soundproof booth using a high quality microphone and stored directly onto a computer at a sample rate of 16 kHz. With the speech editor Praat, the first part of each recorded word up to and including the vowel of the second syllable was excised to serve as an auditory prime to the visual targets. Each experimental target word (e.g., *daffodil*) had an Identity prime, taken from the same word (*daffo* from *daffodil*), a Mismatch prime, taken from the other word of the pair (*defi* from *deficit*), and a Control prime, taken from the unrelated word (*moni* from *monitor*). Note that the Identity prime for one word (*daffo* for *daffodil*) served as the Mismatch prime for the other word of a pair (*daffo* for *deficit*). The experimental target words and their primes are listed in Appendix A.

Further, 24 filler words and 32 filler non-words with Identity primes, and the same number of words and non-words with Mismatch primes and with Control primes were selected and constructed as described for the experimental items. Mismatch primes differed from the visual targets in one vowel, but never in /æ/ or /ɛ/. All primes, including those for non-word targets, were the beginning of a real word. Items selected for visual presentation were not spelled like existing Dutch words, and items selected for auditory presentation did not sound like existing Dutch words.

For the phonetic categorization task, the 24 pairs of experimental Identity and Mismatch primes were used. Note that no continuum with ambiguous tokens was used (cf. Eimas, Marcovitz Hornstein, & Payton, 1990; McQueen, Norris, & Cutler, 1999).

Design

For the cross-modal priming task, the target items were divided into six lists (2 words per pair × 3 conditions). There were 24 pairs of experimental visual target words. Each participant saw only one word of each pair, 12 with an /æ/ and 12 with an /ɛ/. Each participant was presented with eight of the experimental visual targets in each of the three conditions: Identity condition (preceded by auditory presentation of the first two syllables of the same word), Mismatch condition (preceded by the first two syllables of the paired word which overlapped with the first two syllables of the target word, except that an /æ/ in the target was an /ɛ/ in the prime and vice versa), and Control condition (preceded by the first two syllables of the unrelated word). Each participant was presented with all of the filler words and filler non-words, so that each participant saw a total of 96 words and 96 non-words, with 64 presentations in each of the three conditions. Items were presented in a semi-random order, such that maximally five visually presented words or five visually presented non-words occurred in succession, and two experimental targets were separated by at least one other item.

In the phonetic categorization task, participants were presented with four repetitions of the 48 stimuli which served as Identity and Mismatch primes in the previous task. The items were semi-randomized such that the same phoneme occurred maximally five times in succession, and minimally two other stimuli separated the two items of one pair.

Procedure

Participants were tested one at a time in a quiet room. All participants did both the cross-modal priming task and the phonetic categorization task, with a short break in between.

First, for the cross-modal priming task, the participants received written instructions in their native language, informing them that on each trial they would hear part of an English word, directly after which an English word or non-word would appear on a computer screen. They were asked to press a green response button, labeled “yes”, with their dominant hand if they thought the visually presented item was an English word, and a red response button, labeled “no”, with their non-dominant hand if they thought the visually presented item was not an English word. Participants were asked to respond both as fast and as accurately as possible. The task started with 12 practice trials and was controlled with NESU (Nijmegen Experiment Set-Up) software. On each trial, an auditory stimulus was presented and at offset of that, a visual stimulus was presented. The auditory materials were presented binaurally over closed headphones at a comfortable listening level and the visual materials appeared in large font on a computer screen in front of the participants. No time limit was imposed for the responses. After each button press, the next trial started.

After having finished the cross-modal priming task, participants received written instructions for the phonetic categorization task. They were informed that they would hear parts of words containing either an /æ/ or an /ɛ/. They were instructed to decide which of these two sounds they had heard, and to press a green response button, labeled “E”, with their dominant hand when they had heard an /ɛ/ and a red response button, labeled “A”, with their non-dominant hand when they had heard an /æ/. The participants were asked to respond both as fast and as accurately as possible. Before the task started, the participants heard some examples of non-words containing /æ/ or /ɛ/ while the corresponding label (“A” or “E”) appeared on a computer screen, to make it clear, particularly to the Dutch participants, which sounds were intended. The task started with 8 practice trials and was controlled with NESU software. Stimuli were presented binaurally over closed headphones at a comfortable listening level, one at a time. No time limit was imposed for the responses. After each button press, presentation of the next item started.

Results

In this experiment and the following, reaction times (RTs) were measured from item offset, outliers were removed, the proportions of correct responses were arcsine transformed prior to analysis, and RT analyses were performed on the logarithms of the RTs of the correct responses. In the present experiment, the results of one experimental pair had to be excluded due to an error in the item lists.

Cross-modal priming

The hypothesis being tested was that hearing the first part of a word would cause more activation of a word mismatching in the /æ/-/ɛ/ contrast for the nonnative listeners than for the native listeners. Mismatch primes were predicted to facilitate the recognition of visual targets more for the Dutch listeners than for the English listeners. For the English listeners, less facilitation was expected in the Mismatch condition than in the Identity condition, or possibly no facilitation at all. For the Dutch listeners, the amount of facilitation in the Mismatch condition might be similar to that in the Identity condition. Figure 1 shows that this was exactly the pattern found in the proportion of correct responses.

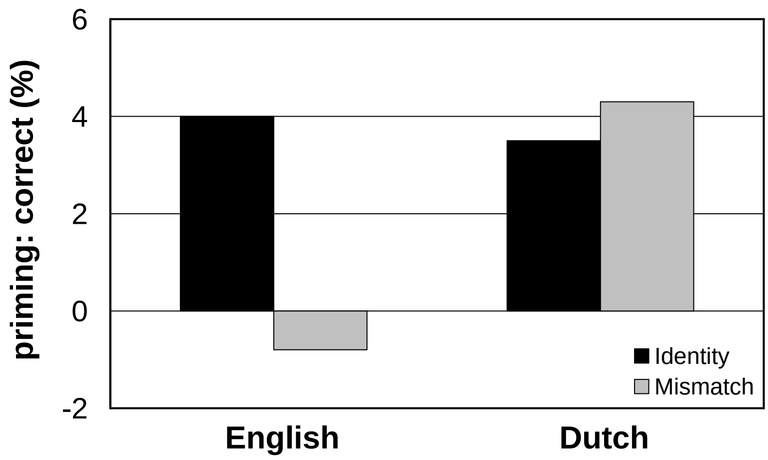

Figure 1. Experiment 1, cross-modal priming: English and Dutch listeners' priming results, computed as the difference between the percentage of correct responses in the Identity or the Mismatch condition and the Control condition, with a positive value indicating facilitation.

Table 1. Experiment 1, cross-modal priming: English and Dutch listeners' percentage of correct responses and RTs of correct responses for target words in Control, Identity, and Mismatch condition, separately for target words containing /æ/ or /ɛ/. Examples are given in brackets.

		correct (%)		RT (ms)	
Target word	**Condition (prime)**	**English**	**Dutch**	**English**	**Dutch**
/æ/ *(daffodil)*	**Control *(moni)***	95.8	70.2	682	828
	Identity *(daffo)*	99.1	74.7	680	751
	Mismatch *(defi)*	92.9	73.5	684	771
/ɛ/ *(deficit)*	**Control *(moni)***	88.0	70.3	729	782
	Identity *(defi)*	92.6	72.8	692	786
	Mismatch *(daffo)*	89.1	75.4	692	801

Proportion correct

Table 1 shows the percentage of correct responses and the RTs of the correct responses. First, the proportions of correct responses were analyzed.

The interaction between native language and condition was significant by subjects but not by items (*F1* (2, 284) = 4.00, $p < .05$; *F2* (2, 90) = 1.52, $p > .1$). For the conditions Identity versus Control, there was no interaction between native language and condition (*F1* (1, 142) = 1.90, $p > .1$; *F2* (1, 45) < 1) and there were more correct responses in the Identity condition than in the Control condition (*F1* (1, 142) = 4.71, $p < .05$; *F2* (1, 45) = 20.46, $p < .001$). For the conditions Mismatch versus Control, there was no interaction between native language and condition (*F1* (1, 142) = 1.63, $p > .1$; *F2* (1, 45) < 1) and no main effect of condition (*F1* (1, 142) < 1; *F2* (1, 45) < 1).

The crucial comparison was between the Identity and the Mismatch conditions. As expected, for these conditions, there was an interaction between native language and condition (*F1* (1, 142) = 10.96, $p < .001$; *F2* (1, 45) = 4.38, $p < .05$). For the Dutch listeners, there was no difference between the Identity and the Mismatch condition (*F1* (1, 71) = 1.23, $p > .1$; *F2* (1, 45) < 1). For the English listeners on the other hand there were more correct responses in the Identity condition than in the Mismatch condition (*F1* (1, 71) = 16.90, $p < .001$; *F2* (1, 45) =12.44, $p < .001$).

A main effect of phoneme was significant in the analysis by subjects but not by items (*F1* (1, 138) = 7.46, $p < .01$; *F2* (1, 44) < 1) and there were no interactions involving phoneme. Thus, the results were similar for words with an /æ/ and words with an /ɛ/.

Overall, the English listeners gave more correct responses than the Dutch listeners (*F1* (1, 142) = 130.55, $p < .001$; *F2* (1, 45) = 33.43, $p < .001$).

Reaction time

In the analysis of the RTs of the correct responses, there were no interactions between native language and condition. There was a main effect of condition (*F1* (2, 284) = 5.54, $p < .01$; *F2* (2, 86) = 4.66, $p < .05$), but pairwise comparisons of the three conditions did not yield significant differences. RTs were shorter for the English listeners than for the Dutch listeners (*F1* (1, 142) = 21.11, $p < .001$; *F2* (1, 43) = 63.07, $p < .001$). There was no main effect of phoneme (*F1* (1, 125) = 2.17, $p > .1$; *F2* (1, 42) < 1), and there were no interactions involving phoneme.

Note that the analysis of the RTs of the correct responses was performed with a considerably reduced data set, due to the large proportion of errors made by the Dutch listeners. This may explain why there were no interactions involving native language, as has been observed in the analysis of the proportion correct.

Phonetic categorization

The phonetic categorization task was included to assess the Dutch listeners' categorization of the /æ/-/ɛ/ contrast, and to compare their performance with that of the English listeners. It was expected that the Dutch listeners would categorize the phonemes less accurately than the English listeners, in line with Broersma (2005).

The results were as expected. The percentage of correct responses to items containing an /æ/ was 85.6 % for the English listeners and 54.6 % for the Dutch listeners, and the percentages correct for items containing an /ɛ/ were 91.7 % and 66.9 %, respectively. The Dutch listeners made more errors than the English listeners did ($F1$ (1, 142) = 938.66, p <.001; $F2$ (1, 44) = 141.95, p < .001).

Further, there were more correct responses for the items containing an /ɛ/ than for the items containing an /æ/ ($F1$ (1, 142) = 57.68, p < .001; $F2$ (1, 44) = 5.25, p < .05). Thus, there was a bias towards perception of /ɛ/.

The Dutch listeners' proportion of correct responses was significantly above chance (50 % correct) (t (71) = 14.31, p < .001 by subjects; t (22) = 7.42, p < .001 by items). However, when the responses to /æ/ and /ɛ/ were assessed separately, the Dutch listeners' proportion of correct responses was significantly above chance for /ɛ/ (t (71) = 13.37, p < .001 by subjects; t (22) = 3.87, p < .001 by items) but not for /æ/ (t (71) = 3.75, p < .001 by subjects; t (22) = 1.38, p > .1 by items).

Discussion

As expected, there was more activation of competitors with partially overlapping onsets for the Dutch listeners than for the English listeners. In the cross-modal priming task, the presentation of Mismatch primes had different effects for the two groups of listeners. Both for native and for nonnative listeners, presentation of an Identity prime facilitated the recognition of the target word. For the English listeners, Mismatch primes did not facilitate the recognition of the target words. For the Dutch listeners on the other hand, Mismatch primes facilitated the recognition of the target words, resulting in as many correct responses in the Mismatch condition as in the Identity condition.

The results from the categorization task showed that, as predicted, the Dutch listeners recognized the vowels less accurately than the English listeners did, in line with Broersma (2005). Further, there was a bias towards perception of /ɛ/ for Dutch and English listeners alike. The Dutch listeners categorized the items with an /ɛ/ but not those with an /æ/ with a level of accuracy above chance. The Dutch listeners' low level of perceptual accuracy can explain the finding that presentation of the onset of words activated competitors with partially overlapping onsets more for Dutch than for English listeners.

Experiment 2

Experiment 2 investigated Dutch and English listeners' perception of English minimal pairs. Apart from the occurrence of increased lexical activation in nonnative listening, it also assessed the source of such an increase.

As discussed above, increased competitor activation may result from perceptual ambiguity. However, Broersma (2002, submitted) also found increased competitor activation for non-words which were not based on a difficult to distinguish contrast. Whereas some non-words were based on the difficult to distinguish /æ/-/ɛ/ contrast (e.g., *lamp* – *lemp*), others differed from real words in contrasts that were easy to distinguish for the nonnative listeners, namely in the voicing of the word-final consonant (e.g., *globe* – *glope*). Non-words based on easy to distinguish contrasts caused more lexical activation for the Dutch listeners than for native listeners, just like the non-words based on the perceptually ambiguous contrast did. Thus, perceptual ambiguity does not seem to be the only possible cause of increased lexical activation in the perception of nonnative speech.

The present study investigated the perception of minimal pairs based on easy or on difficult to distinguish phoneme contrasts. Some of the minimal pairs differed in the /æ/-/ɛ/ contrast, which is difficult to distinguish for Dutch listeners (Broersma, 2005). Other minimal pairs were based on the word-final consonant voicing contrasts /z/-/s/, /b/-/p/, and /d/-/t/. These phonemes exist in Dutch and are very similar to the British English phonemes. As Dutch does not allow for voiced obstruents at the end of words in isolation, the voicing contrast does not occur in word-final position in Dutch. Cutler, Weber, Smits, and Cooper (2004) found that Dutch listeners identified syllable-final voicing distinctions in American English consonants less accurately than native listeners did. However, studies that focused on the perception of obstruents which closely match Dutch phonemes (including those used in the present study) showed that Dutch listeners accurately distinguished the voicing of these contrasts in syllable-final position. Van Wieringen (1995) presented listeners with voiced and voiceless consonants pronounced by a Dutch speaker and found that Dutch listeners categorized final voiced and voiceless plosives equally accurately, and as accurately as native English listeners did. Further, Dutch listeners were found to categorize British English word-final obstruent voicing contrasts as accurately as native listeners of English did when the contrasts occurred at the end of clear non-words (Broersma, 2005) or in words and non-words in which the voiced consonant of a real word was replaced with a voiceless consonant and vice versa (Broersma, submitted). Therefore, in the present experiment, if there was more competitor activation for nonnative listeners than for native listeners for minimal pairs based on the word-final voicing contrasts, this could not be attributed to perceptual ambiguity.

The experiment again consisted of a cross-modal priming task and a phonetic categorization task. The phonetic categorization task was included to verify that the vowel contrast but not the consonant voicing contrasts was perceptually ambiguous for the Dutch listeners.

Method

Participants

Different subjects participated in the two tasks. In the cross-modal priming task 72 native speakers of Dutch and 72 native speakers of British English participated. In the phonetic categorization task 20 native speakers of Dutch and 20 native speakers of British English participated. The participants met the description given for Experiment 1, except that for the cross-modal priming task the English participants were recruited at the University of Birmingham. None had participated in Experiment 1.

Materials

For the cross-modal priming task, 21 pairs of English words were selected as visual target words. The words of a pair were identical, except for one phoneme. For six pairs, one word contained an /æ/ and the other an /ɛ/ (e.g., *flash-flesh*). For 15 pairs, one word contained a voiced final consonant and the other a voiceless final consonant (e.g., *robe-rope*). Four of the word pairs based on the /æ/-/ɛ/ contrast were disyllabic and the others monosyllabic. The mean logarithmic lemma frequency per million, determined with the CELEX lexical database (Baayen et al., 1995), was 1.32 for words with an /æ/, 1.65 for words with an /ɛ/, 1.66 for words with a voiced final consonant, and 1.93 for words with a voiceless final consonant. For each pair, a phonologically and semantically unrelated word was selected with the same number of syllables as the target words.

All words were recorded by the same speaker and in the same manner as described for Experiment 1. Each experimental visual target word (e.g., *flash*) had an auditory Identity prime (the same word, *flash*), a Mismatch prime (the other word of the pair, *flesh*), and a Control prime (the unrelated word, *spite*). The Identity prime for one word served as the Mismatch prime for the other word of a pair. The experimental target words and their primes are listed in Appendix B.

Further, 21 filler words with Identity primes, 21 with Mismatch primes, and 21 with Control primes, as well as 42 filler non-words with Mismatch primes and 42 with Control primes were selected and constructed as described for the experimental items. Mismatch primes differed from the visual targets either in the vowel or in the final consonant (but never in the /æ/-/ɛ/ contrast or a final consonant voicing contrast), proportional to the number of experimental items. The number of mono- and disyllabic items was also proportional to the experimental items. All primes, including those for non-word targets, were real words. For that reason, there were no Identity primes for non-words. Items selected for visual presentation were not spelled like existing Dutch words, and items selected for auditory presentation did not sound like existing Dutch words.

For the phonetic categorization task, the 21 pairs of experimental Identity and Mismatch primes were used.

Design

For the cross-modal priming task, the target items were divided into six lists (2 words per pair × 3 conditions), with vowel items and consonant items distributed evenly over the lists, and the target phonemes distributed as evenly as possible. There were 21 pairs of experimental visual target words. Each participant saw only one word of each pair. Each participant was presented with seven of the experimental visual targets in each of the three conditions: Identity condition (preceded by auditory presentation of the same word), Mismatch condition (preceded by the paired word which mismatched with the target in one phoneme), and Control condition (preceded by the unrelated word). Each participant was presented with all of the filler words and filler non-words, so that each participant saw a total of 84 words and 84 non-words. Items were presented in a random order.

For the phonetic categorization task, the 42 stimuli which served as Identity and Mismatch primes in the cross-modal priming task were used. The items were presented in four blocks, each block testing one phoneme contrast. The contrasts were /æ/-/ɛ/, /z/-/s/, /b/-/p/, and /d/-/t/. The /æ/-/ɛ/ block contained 12 target words, the /z/-/s/ block 6, the /b/-/p/ block 4, and the /d/-/t/ block 20. Each block also contained a varying number of filler words and non-words, with equal numbers of the two target phonemes in each block. All items were presented four times, semi-randomized as in the phonetic categorization task of Experiment 1.

Procedure

For the cross-modal priming task, the procedure was as described for Experiment 1, except that the participants were instructed that they would hear a word (as opposed to part of a word as in Experiment 1).

For the phonetic categorization experiment, the procedure was largely as described for Experiment 1. However, four contrasts were tested in the present experiment. The participants were first instructed about the procedure (as in Experiment 1). Before each block, they received further information about the two response alternatives in that block, and about the position of the target phoneme. Before the /æ/-/ɛ/ block, participants heard some examples of these phonemes (as in Experiment 1). For the other blocks, it was assumed that the participants understood without further illustration which phonemes were intended. The practice part for the vowel block contained 12 trials and those for the consonant blocks 8. The response buttons were labeled “A” and “E”, “Z” and “S”, “B” and “P”, or “D” and “T”, respectively, where the first one of each pair was to be pressed with the non-dominant hand and the second with the dominant hand.

Results

Cross-modal priming

Unlike in Experiment 1, auditory primes consisted of entire words. Lexical representations are expected to receive more activation upon presentation of a full word than upon presentation of an incomplete word (Gaskell & Marslen-Wilson, 2002). After presentation of a full word, lexical representations might receive enough activation to exert strong inhibitory effects on other active representations. Therefore, inhibitory as well as facilitatory effects were expected. For the English listeners, Identity primes were predicted to facilitate recognition of the visual targets, and Mismatch primes were predicted to inhibit it.

The hypothesis being tested was that presentation of one word would cause more activation of its minimal pair for the nonnative listeners than for the native listeners. Therefore, for the Dutch listeners, less inhibition was expected in the Mismatch condition than for the English listeners, or possibly no inhibition at all. Figure 2 shows that the RTs of the correct responses were as predicted, with inhibition in the Mismatch condition for the English listeners, but not for the Dutch listeners.

Neither in the analysis of the proportion of correct responses, nor in the analysis of the RTs of the correct responses was there an interaction involving item type (vowel items vs. consonant items). When vowel items and consonant items were analyzed separately, there were no interactions with or main effects of phoneme (/æ/ vs. /ɛ/) or consonant voicing. Therefore, vowel items and consonant items were analyzed together.

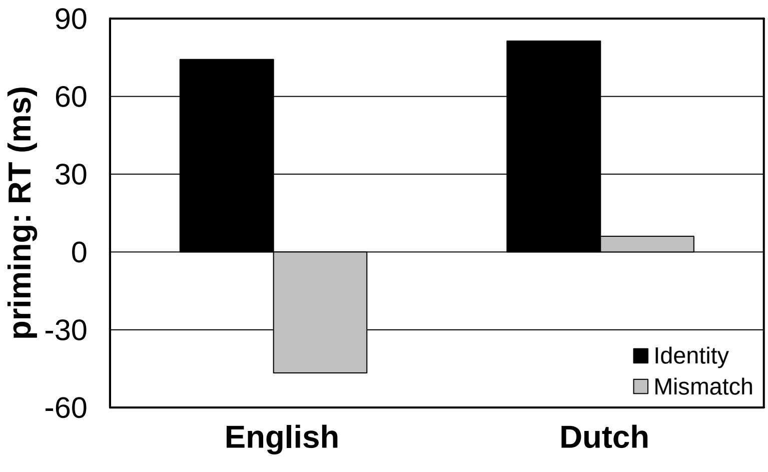

Figure 2. Experiment 2, cross-modal priming: English and Dutch listeners' priming results, computed as the difference between the reaction times of correct responses in the Identity or the Mismatch condition and the Control condition, with a positive value indicating facilitation.

Table 2. Experiment 2, cross-modal priming: English and Dutch listeners' percentage of correct responses and RTs of correct responses for target words in Control, Identity, and Mismatch condition, separately for target words containing /æ/ or /ɛ/, or a voiced or voiceless final consonant. Examples are given in brackets.

		correct (%)		RT (ms)	
Target word	**Condition (prime)**	**English**	**Dutch**	**English**	**Dutch**
/æ/ *(flash)*	**Control *(spite)***	95.8	76.4	646	743
	Identity *(flash)*	95.8	79.2	538	718
	Mismatch *(flesh)*	81.9	64.8	728	764
/ɛ/ *(flesh)*	**Control *(spite)***	100	90.1	632	738
	Identity *(flesh)*	97.2	88.9	569	680
	Mismatch *(flash)*	90.3	75.0	668	731
Voiced *(robe)*	**Control *(suck)***	94.4	92.2	620	740
	Identity *(robe)*	99.4	93.3	538	637
	Mismatch *(rope)*	91.1	81.7	681	734
Voiceless *(rope)*	**Control *(suck)***	97.8	92.7	613	701
	Identity *(rope)*	97.8	93.3	556	612
	Mismatch *(robe)*	93.3	89.3	638	689

Proportion correct

Table 2 shows the percentage of correct responses and the RTs of the correct responses. There was no interaction between native language and condition (*F1* (2, 284) < 1; *F2* (2, 38) < 1). There was a main effect of condition (*F1* (2, 284) = 23.95, *p* < .001; *F2* (2, 38) = 23.58, *p* < .001). There was no difference between the Identity and the Control condition (*F1* (1, 142) < 1; *F2* (1, 19) <1). There were more errors in the Mismatch condition than in the Control condition (1, 142) = 27.35, *p* < .001; *F2* (1, 19) = 37.55, *p* < .001) and there were more errors in the Mismatch condition than in the Identity condition (*F1* (1, 142) = 33.54, *p* < .001; *F2* (1, 19) = 37.97, *p* < .001). Thus, there was inhibition in the Mismatch condition, for Dutch and English listeners alike.

Further, there were more correct responses to consonant items than to vowel items (*F1* (1, 142) = 9.56, *p* < .01; *F2* (1, 19) = 6.62, *p* < .05). This may be explained by the frequency of the target words, which was higher for the consonant items than for the vowel items. The Dutch listeners made more errors than the English listeners did (*F1* (1, 142) = 48.59, *p* < .001; *F2* (1, 19) = 25.20, *p* < .001).

Reaction time

Crucially, in the analysis of the RTs of the correct responses, there was an interaction between native language and condition (*F1* (2, 256) = 4.41, *p* < .05; *F2* (2, 38) = 8.98, *p* < .001).

For the conditions Identity versus Control, there was no interaction with native language (*F1* (1, 137) < 1; *F2* (1, 19) <1), and RTs were shorter in the Identity than in the Control condition (*F1* (1, 137) = 75.32, $p < .001$; *F2* (1, 19) = 33.88, $p < .001$).

For the conditions Mismatch versus Control, there was an interaction between native language and condition (*F1* (1, 129) = 5.27, $p < .05$; *F2* (1, 19) = 10.72, $p < .01$). For the English listeners, RTs were longer in the Mismatch condition than in the Control condition (*F1* (1, 69) = 12.55, $p < .001$; *F2* (1, 19) = 11.18, $p < .01$). For the Dutch listeners, on the other hand, there was no difference between these conditions (*F1* (1, 60) < 1; *F2* (1, 19) < 1).

For the conditions Identity versus Mismatch, there was also an interaction between native language and condition (*F1* (1, 132) = 4.63, $p < .05$; *F2* (1, 19) = 12.42, $p < .01$). RTs were longer in the Mismatch condition than in the Identity condition, both for the English listeners (*F1* (1, 69) = 97.15, $p < .001$; *F2* (1, 19) = 109.38, $p < .001$) and for the Dutch listeners (*F1* (1, 63) = 16.60, $p < .001$; *F2* (1, 19) = 12.76, $p < .01$). However, the difference was larger for the English listeners than for the Dutch listeners.

Further, RTs were shorter for the English listeners than for the Dutch listeners (*F1* (1, 128) = 13.32, $p < .001$; *F2* (1, 19) = 142.58, $p < .001$).

Phonetic categorization

Table 3 shows the English and Dutch listeners' percentage of correct responses. The phonetic categorization task assessed to which extent the Dutch listeners could distinguish the /æ/-/ɛ/ contrast and the consonant voicing contrast. The prediction was that the Dutch listeners would categorize the /æ/-/ɛ/ contrast less accurately than the English listeners (as in Experiment 1 and in line with Broersma, 2005) and the consonant voicing contrast as accurately as the English listeners (in line with Broersma, 2005).

For the vowel items, as predicted, the Dutch listeners made more errors than the English listeners did (*F1* (1, 38) = 101.61, $p < .001$; *F2* (1, 5) = 142.86, $p < .001$). There were more correct responses for the items containing /ɛ/ than for the items containing /æ/ (*F1* (1, 38) = 30.77, $p < .001$; *F2* (1, 5) = 8.05, $p < .05$). Thus, there was a bias towards perception of /ɛ/, for English and Dutch listeners alike. The Dutch listeners' proportion of correct responses was significantly above chance level (t (19) = 3.82, $p < .001$ by subjects; t (5) = 4.55, $p < .01$ by items). However, when the items with /æ/ and /ɛ/ were assessed separately, the difference from chance level was no longer significant, neither for /æ/ (t (19) < 1 by subjects; t (5) < 1 by items) or for /ɛ/ (t (19) = 4.65 , $p < .001$ by subjects; t (5) = 2.05, $p = .096$ by items).

For the consonant items, again as predicted, there was no difference between the English and the Dutch listeners' proportions of correct responses (*F1* (1, 38) < 1; (*F2* (1, 14) < 1). There was no difference between voiced and voiceless consonants (*F1* (1, 38) = 3.11, $p = .086$; *F2* (1, 14) <1).

Table 3. Experiment 2, phonetic categorization: English and Dutch listeners' percentage of correct responses, separately for items containing /æ/ or /ɛ/, or a voiced or voiceless final consonant. Examples are given in brackets.

Stimulus	English	Dutch
/æ/ *(flash)*	82.1	51.4
/ɛ/ *(flesh)*	97.1	69.0
Voiced *(robe)*	90.8	94.1
Voiceless *(rope)*	92.6	89.5

Discussion

As predicted, in the cross-modal priming task, presentation of one word caused more activation of its minimal pair for the nonnative listeners than for the native listeners. After an Identity prime, there was facilitation in the RTs for the English and the Dutch listeners alike. After a Mismatch prime, there was inhibition in the RTs for the English listeners, and neither facilitation nor inhibition for the Dutch listeners.

The Dutch listeners did not always treat the minimal pairs as homophones, with presentation of either word activating both lexical representations, as this should have resulted in similar RTs in the Identity and the Mismatch conditions. The results suggest that presentation of one word sometimes caused activation of the identical word and inhibition of the paired word, whereas in other cases both words remained activated and neither one inhibited the other. Thus, in the Identity condition there would always be facilitation, and in the Mismatch condition, there would be inhibition in some cases and facilitation in other cases.

For the vowel items, this can be explained by the finding from the phonetic categorization task that the minimal pairs were perceptually ambiguous for the Dutch listeners. Connine et al. (1994) found that ambiguous stimuli which were equally consistent with two interpretations activated both lexical representations to the same extent. In the present experiment, the stimuli were not fully ambiguous to the Dutch listeners, as these listeners categorized the vowels at a level of accuracy which was just above chance level. This is in line with the finding that the stimuli sometimes activated both lexical representations to the same extent, whereas in other cases the difference in activation of the two representations was enough for the target word to inhibit its competitor.

For the consonant items, the results from the cross-modal priming task cannot be attributed to perceptual ambiguity. The phonetic categorization task showed that the Dutch listeners recognized the consonant voicing contrasts as accurately as the English listeners did. Nevertheless, both words of a minimal pair remained activated after presentation of one of them as often as for the vowel items. A possible explanation is that the Dutch listeners sometimes disregarded the voicing distinction for lexical access, even though they could accurately distinguish the contrast. As voiced obstruents do not occur at the end of words in isolation in Dutch, voicing is never distinctive word-finally. This may have induced the Dutch listeners to disregard word-final voicing in English as well. Another possible explanation has to do with the Dutch listeners' lexical representations of minimal pairs. As Dutch speakers may not always pronounce the voicing of word-final obstruents in English correctly, Dutch listeners may incorporate this in their lexical representations. Thus, some minimal pairs may have been stored as homophones with similar form representations. For these pairs, presentation of one word could activate both lexical representations to a similar extent.

Note that there was no evidence for a bimodal distribution. It was not the case that some of the nonnative listeners processed all minimal pairs as homophones whereas other listeners processed them all in a native-like manner. Neither was it the case that the nonnative listeners always processed certain minimal pairs as homophones and other minimal pairs in a native-like manner. Rather, the data showed that for all the nonnative listeners and for all the minimal pairs, primes sometimes activated the competitor word.

General discussion

The results of this study show that partially overlapping words cause more competitor activation for nonnative listeners than for native listeners. This was found both for words with partially overlapping onsets and for minimal pairs. Similar results were found for minimal pairs based on contrasts which were difficult or easy to distinguish for the nonnative listeners. Whether a contrast was difficult or easy to distinguish was corroborated with phonetic categorization tasks.

For the vowel contrast, both words with partially overlapping competitors and minimal pairs were found to cause more competitor activation for nonnative listeners than for native listeners. This increased activation of lexical competitors based on the vowel contrast can be explained by the perceptual ambiguity of these vowels for the Dutch listeners. The finding that the vowels /æ/ and /ɛ/ were not fully ambiguous for the Dutch listeners may seem discrepant with the finding in Experiment 1 that there was no difference in the proportion of correct responses for these listeners between the Identity and the Mismatch condition. However, the RTs, which might reflect more subtle differences in lexical activation, were numerically (although not statistically significantly) shorter in the Identity condition than in the Mismatch condition. Further, the Dutch listeners' ability to distinguish the vowels at a level of accuracy above chance is in line with the finding in Experiment 2 that one word of a minimal pair sometimes but not always activated both lexical representations to the same extent.

For the consonant voicing contrast, minimal pairs were also found to cause more competitor activation for nonnative listeners than for native listeners. Presentation of one word of a minimal pair sometimes led to the activation of both words to a similar extent, even though the consonant contrasts were not perceptually ambiguous for the Dutch listeners. This activation of both lexical representations occurred as often as for the minimal pairs based on the difficult to distinguish /æ/-/ɛ/ contrast. Broersma (submitted) also found that English non-words differing from words in the final consonant voicing contrast caused more lexical activation for Dutch listeners than for native English listeners. Thus, increased lexical activation is not restricted to words containing phoneme contrasts that are difficult to distinguish for nonnative listeners, but also occurs for words containing easy to distinguish contrasts. Therefore, the increased activation of lexical competitors may be more pervasive than could be suspected on the basis of previous research.

The two phonetic categorization tasks confirmed that the Dutch listeners recognized the vowels /æ/ and /ɛ/ with a level of accuracy just above chance, which seemed to be due to their categorization of /ɛ/ rather than their categorization of /æ/. They categorized the vowels less accurately than the English listeners did. Both the English and the Dutch listeners had a bias towards perception of /ɛ/. These results are consistent with previous studies. Schouten (1975) found that the contrast was difficult to distinguish for Dutch listeners. The finding that the Dutch listeners distinguished the contrast at a level of accuracy above chance but less accurately than the English listeners was similar to the results of a phonetic categorization experiment with clear non-words (Broersma, 2005), and to the results of three phonetic categorization experiments with real words and non-words in which the /æ/ of a real word was replaced with an /ɛ/ and vice versa (Broersma, submitted), all with Dutch listeners from the same population as in the present study. A bias towards perception of /ɛ/ was also found for English and Dutch listeners alike in the three categorization experiments of Broersma (submitted), who proposed that this bias might be the result of the frequency of occurrence in English, which is higher for /ɛ/ than for /æ/. The phonetic categorization task of Experiment 2 showed that the Dutch listeners categorized the word-final consonant voicing contrasts as accurately as the English listeners did. This is also in line with the results from previous studies (Broersma, 2005, submitted).

Cutler (2005) computed the upper bounds of the effects of perceptual ambiguity on the activation of lexical competitors. Lexical statistics were computed to determine the potential number of competitors added by perceptual ambiguity of the /æ/-/ɛ/ contrast in English. If the /æ/-/ɛ/ contrast was perceptually fully ambiguous, the number of non-words which occurred embedded in other words and which might be perceived as real words would be considerable, with more than 78,000 occurrences per million words. The number of minimal pairs that would be perceived as homophones was relatively small, with 137 cases. The number of temporarily overlapping competitors, on the other hand, was very large, with an average of 274 added competitors per word.

For the Dutch listeners in the present study, the /æ/-/ɛ/ contrast was not fully ambiguous, and the number of added lexical competitors due to misperception of this contrast is likely to be smaller than the maximum that Cutler (2005) computed. However, the statistics given by Cutler (2005) indicate that the possible effect of one ambiguous contrast only is already considerable. Of course listeners may be confronted with many perceptually ambiguous contrasts while listening to a second language. Further, the results of the present study show that increased activation of lexical competitors occurred not only after presentation of words that contained difficult to distinguish phoneme contrasts, but also for words containing easy to distinguish contrasts. Processing of all of these contrasts may simultaneously increase the number of activated lexical representations. Further, the number of possible lexical competitors may increase sharply due to the combination of several of these contrasts within a single word. For example, presentation of the word *bad* may not only activate the lexical representation of *bed* for Dutch listeners, but it may activate *bat* and *bet* as well. Thus, the increase of lexical activation may be very large in nonnative listening.

An increase of lexical activation can influence speech recognition in several ways. In the case of minimal pairs, lexical competitors may receive more activation in nonnative listening than in native listening. In the most problematic cases, nonnative listeners may not even be able to determine on the basis of the phonetic input which word they have heard, and they may have to rely on the context to select one interpretation. Thus, while such minimal pairs may be relatively rare (Cutler, 2005), they may slow down the recognition process considerably. In the case of partially overlapping words, at some point the speech input will clearly deviate from one of the lexical representations, which will solve the competition. However, the competitor may remain active much longer for nonnative listeners than for native listeners.

As it is more difficult to recognize a word when more lexical competitors are active, an increase in lexical activation is harmful to speech recognition. Although the activation of lexical competitors is a necessary part of speech comprehension (see e.g., McQueen, 2004), it has also been found to complicate the recognition of spoken words for native listeners (Luce et al., 1990; Norris et al., 1995; Vroomen & De Gelder, 1995). An increase in lexical activation extends this problem for nonnative listeners. The results from the present study show that the increase of competitor activation in nonnative listening may be very large and may seriously complicate the recognition of speech in a second language.

References

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

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.

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

Broersma, M. (2005). Perception of familiar contrasts in unfamiliar positions. *Journal of the Acoustical Society of America, 117*, 3890-3901. *Chapter 2 of this dissertation.*

Broersma, M. (submitted). Spurious lexical activation in nonnative listening. *Cognition. Chapter 3 of this dissertation*.

Connine, C. M., Blasko, D. G., & Wang, J. (1994). Vertical similarity in spoken word recognition: Multiple lexical activation, individual differences, and the role of sentence context. *Perception and Psychophysics, 56*, 624-636.

Cutler, A. (2005). *The lexical statistics of word recognition problems caused by L2 phonetic confusion*. Proceedings of the 9th European Conference on Speech Communication and Technology, Lisbon, September 2005.

Cutler, A., & Otake, T. (2004). *Pseudo-homophony in non-native listening*. Poster presented at the 75th meeting of the Acoustical Society of America, New York.

Cutler, A., Weber, A., & Otake, T. (in press). Asymmetric mapping from phonetic to lexical representations in second-language listening. *Journal of Phonetics*.

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.

Eimas, P. D., Marcovitz Hornstein, S. B., & Payton, P. (1990). Attention and the role of dual codes in phoneme monitoring. *Journal of Memory and Language, 29*, 160-180.

Gaskell, M. G., & Marslen-Wilson, W. D. (2002). Representation and competition in the perception of spoken words. *Cognitive Psychology, 45*, 220-266.

Luce, P. A., Pisoni, D. B., & Goldinger, S. D. (1990). Similarity neighborhoods of spoken words. In G. T. M. Altmann (Ed.), *Cognitive Models of Speech Processing:*

Psycholinguistic and Computational Perspectives (pp. 122-147). Cambridge, MA: MIT.

McQueen, J. (2004). Speech perception. In K. Lamberts & R. Goldstone (Eds.), *The Handbook of Cognition* (pp. 255-275). London: Sage Publications.

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.

McQueen, J. M., Cutler, A., Briscoe, T., & Norris, D. (1995). Models of continuous speech recognition and the contents of the vocabulary. *Language and Cognitive Processes, 10*, 309-331.

McQueen, J. M., Norris, D., & Cutler, A. (1999). Lexical influence in phonetic decision making: Evidence from subcategorical mismatches. *Journal of Experimental Psychology: Human Perception and Performance, 25*, 1363-1389.

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.

Onifer, W., & Swinney, D. A. (1981). Accessing lexical ambiguities during sentence comprehension: Effects of frequency of meaning and contextual bias. *Memory and Cognition, 9*, 225-236.

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.

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.

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.

Van Wieringen, A. (1995). *Perceiving Dynamic Speechlike Sounds: Psycho-acoustics and Speech Perception.* Unpublished doctoral dissertation, Universiteit van Amsterdam.

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.

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

Zwitserlood, P. (1996). Form priming. *Language and Cognitive Processes, 11*, 589-596.

Appendix A

Experimental stimuli used in Experiment 1.

Underlined fragments served as primes. Within each triplet the onset of each target served both as the Identity prime for that target and as the Mismatch prime for the other target.

Target 1	Target 2	Control prime
janitor	genitive	prosody
adjective	educate	cylinder
adequate	editor	permanent
accident	execute	poverty
allergy	eloquent	formula
amorous	emerald	optimist
animal	enemy	property
animate	enervate	vocalise
antelope	entity	sinister
appetite	epilogue	civilise
avalanche	evident	immigrant
banister	benefit	incident
daffodil	deficit	monitor
family	feminine	principle
character	kerosene	article
clarify	clerical	symphony
lavender	levitate	fundament
massacre	messenger	orthodox
parasite	periscope	barbecue
patronise	petrify	fulminate
sacrifice	secretary	realise
salary	celebrate	funeral
tantalise	tentacle	mutiny

Appendix B

Experimental stimuli used in Experiment 2.

Within each triplet each target served both as the Identity prime for that target and as the Mismatch prime for the other target.

Target 1	Target 2	Control prime
flash	flesh	spite
tan	ten	blood
cattle	kettle	fever
mansion	mention	bottle
mantle	mental	passage
marry	merry	dozen
phase	face	home
rise	rice	chief
lose	loose	judge
cab	cap	stiff
robe	rope	suck
bride	bright	shave
code	coat	mouse
fade	fate	pipe
grade	great	home
greed	greet	flush
hard	heart	touch
hide	height	youth
slide	slight	trap
thread	threat	fresh
wide	white	close

Summary and conclusions

Chapter 5

Summary

Phonetic processing

This dissertation investigated several aspects of the perception of speech in a second language. One of its main topics is phonetic processing in a second language. The perception of nonnative phoneme contrasts with different degrees of similarity to the native language phonology was investigated. First, Dutch listeners' perception of the British English /æ/-/ɛ/ contrast was assessed. As Dutch has only one phoneme in the perceptual space of the English /æ/ and /ɛ/, the Perceptual Assimilation Model (Best, 1994; Best, McRoberts, & Sithole, 1988) predicts it to be among the most difficult contrasts for Dutch listeners. The nonnative listeners' perception of the vowel contrast was assessed in much detail, and directly compared to native listeners' perception. Second, this study investigated how listeners perceive nonnative phoneme contrasts which are similar to native contrasts, but in a position where they do not occur in the native language. The consonant voicing contrasts /z/-/s/, /v/-/f/, /b/-/p/, and /d/-/t/ exist in Dutch as well as in English and are phonetically quite similar in both languages. However, in Dutch only the voiceless consonants can occur at the end of words in isolation (Booij, 1995). This study assessed Dutch and English listeners' perception of the English consonant voicing contrasts in final position. Attention was given to the necessity to use perceptual cues in a native-like manner.

Vowels

The perception of the /æ/-/ɛ/ contrast was studied in six phonetic categorization experiments reported in Chapters 2, 3, and 4. In Experiment 2.1, Dutch and English listeners categorized a series of unedited recordings of the non-words *faf* and *fef*, containing an /æ/ or an /ɛ/, respectively. In Experiments 3.1, 3.2, and 3.3, listeners categorized real words (e.g., *lamp*) and near-words (e.g., *lemp*). In Experiment 4.1, they categorized fragments of real words (e.g., *daffo* from *daffodil* or *defi* from *deficit*), and in Experiment 4.2 full word forms (e.g., *flash* or *flesh*). In all these experiments, the Dutch listeners categorized the vowels less accurately than the English listeners did. The only exception in these highly consistent results was that the difference between the Dutch and the English listeners' accuracy did not reach statistical significance in Experiment 3.3. In all experiments, the Dutch listeners categorized the vowels with a level of accuracy above chance. In Experiment 2.1, the level of accuracy was amply above chance, whereas in the other experiments it was just above chance. In those experiments, the difference from chance level seemed to be due to the Dutch listeners' categorization of /ɛ/ rather than /æ/.

The percentage of correct responses in Experiment 2.1 was considerably higher than that in the other phonetic categorization experiments. In Experiment 2.1, the Dutch listeners gave 95 % correct responses and the English listeners 99 %, whereas the mean percentage of correct responses across the other experiments was 59 % for the Dutch listeners and 84 % for the English listeners. One important difference was that in Experiment 2.1 only clear non-words were used, whereas in the other experiments the vowels were presented in near-words and words. Previous research has shown that listeners tend to classify phonemes in such a way that the input is consistent with a real word (e.g., Pitt & Samuel, 1993). Therefore, errors on near-words do not necessarily result from perceptual difficulty. Indeed, in Experiments 3.1, 3.2, and 3.3, strong lexical effects were found. Lexical effects may have led to errors in the phonetic categorization tasks of Experiments 4.1 and 4.2 as well. Further, in Experiment 2.1 listeners heard several tokens of the same two stimuli, whereas the other experiments contained many different stimuli. The similar phonetic contexts may have made the target phonemes easier to recognize in Experiment 2.1.

Another difference between the results from Experiment 2.1 and the other phonetic categorization experiments was that a bias towards perception of /ɛ/ was found in all but Experiment 2.1. In the other phonetic categorization experiments there were more correct responses to items with an /ɛ/ than to items with an /æ/, for Dutch and English listeners alike. The absence of this pattern in Experiment 2.1 may have been due to a ceiling effect. For the Dutch listeners, a possible explanation for the occurrence of the bias could be that in their perception the English /ɛ/ may have matched the Dutch /ɛ/ better than the English /æ/ did. However, this cannot explain why the English listeners had a similar bias. Therefore, a more likely explanation might be that the bias was a result of the frequency of occurrence of the phonemes, which is higher for /ɛ/ than for /æ/ (see Chapter 3).

Consonants

The perception of the consonant voicing contrasts /z/-/s/, /v/-/f/, /b/-/p/, and /d/-/t/ was studied in five phonetic categorization experiments, reported in Chapters 2, 3, and 4. In Experiment 2.1, listeners heard several presentations of two non-words, differing in a consonant voicing contrast (e.g., *zeef* and *seef*). The contrasts occurred at the beginning of the non-word (where they could also occur in Dutch) in one part of the experiment, and at the end (where they could not occur in Dutch) in another part of the experiment. In Experiments 3.1 and 3.3, listeners categorized voiced and voiceless consonants at the end of real words (e.g., *globe*) and near-words (e.g., *glope*) and in Experiment 4.2 at the end of words (e.g., *robe* or *rope*). In all these experiments, the Dutch listeners categorized the final voicing contrasts in non-words and in words as accurately as the English listeners did. (Again, errors on the near-words resulted at least partially from lexical effects instead of perceptual ambiguity.) Categorization of the final consonants was unbiased. In Experiment 2.1, the Dutch listeners categorized the voicing contrasts as accurately in final position as in initial position.

In Experiment 2.2, Dutch and English listeners' use of vowel duration as a cue to the final /z/-/s/ and /v/-/f/ contrasts was investigated. Stimulus materials were constructed such that they discouraged the use of vowel duration as a cue to voicing. Vowel duration was uninformative and mismatched other information in the signal for part of the stimuli. The results showed that the Dutch listeners did not base their responses on vowel duration. English listeners on the other hand relied heavily on vowel duration for their categorization of the final /v/-/f/ contrast. As the English listeners were misled by the uninformative vowel duration, their categorization was less categorical than the Dutch listeners' categorization when the final /v/-/f/ contrast was preceded by a long vowel. Although this does not imply that the Dutch listeners never used vowel duration as a cue to final voicing, it shows that they relied on this cue less consistently than the native listeners did.

Lexical processing

The other main topic investigated in this dissertation is lexical processing in a second language. The occurrence of increased lexical activation in nonnative as compared to native listening was investigated. The lexical processing of different types of speech input was examined. First, the processing of near-words, which differed from a real word in one phoneme, was assessed. Near-words may occur in normal speech embedded in other words. It was investigated whether presentation of near-words caused more activation of real words for the nonnative listeners than for the native listeners. Second, the processing of partially overlapping words was assessed. Word pairs had overlapping onsets or fully overlapped, except for one phoneme. It was investigated whether presentation of one word caused more activation of the lexical competitor for the nonnative listeners than for the native listeners.

Near-words

Chapter 3 investigated whether hearing a near-word (e.g., *lemp*) led to more activation of the nearest lexical representation (*lamp*) for nonnative listeners than for native listeners. Two types of near-words were used. Near-words of the first type were based on the /æ/-/ɛ/ contrast, which was difficult to distinguish for the Dutch listeners. Near-words of the second type were based on the word-final consonant voicing contrasts /z/-/s/, /v/-/f/, /b/-/p/, and /d/-/t/, which are not contrastive in Dutch in that position either, but which the Dutch listeners nevertheless distinguished as accurately as the English listeners did. In Experiment 3.1, the items were recorded in isolation. In Experiment 3.2, they were excised from a carrier word.

For example, *deaf* was excised from *definite*, and *daf* from *daffodil*. In Experiment 3.3, stimuli were excised from a two-word context. For example, *lemp* was excised from *evil empire*.

Consistent evidence was found that near-words caused more lexical activation for the nonnative listeners than for the native listeners. The lexical decision task of Experiment 3.1 showed that upon presentation of a near-word, Dutch listeners thought more often that they had heard a real word than English listeners did. In Experiments 3.1, 3.2, and 3.3, phonetic categorization tasks were used, employing the well-established finding that listeners tend to classify phonemes in such a way that the input is consistent with a real word (e.g., Pitt & Samuel, 1993), which is evidence of lexical activation (McClelland & Elman, 1986; Norris, McQueen, & Cutler, 2000). There was more lexical activation after presentation of near-words for the Dutch listeners than for the English listeners. Other than for the vowel items in Experiment 3.3., this pattern was found consistently in the phonetic categorization tasks of Experiments 3.1, 3.2, and, for the consonant items, 3.3. The cross-modal priming tasks of Experiments 3.2 and 3.3 also showed that near-words caused more lexical activation for nonnative than for native listeners.

Further, the cross-modal priming tasks of Experiments 3.2 and 3.3 showed that for the Dutch listeners, near-words often caused as much lexical activation as real words did. First, this was the case for the near-words based on the vowel contrast. Second, for the near-words based on the consonant contrast, there was a difference between the near-words with a voiced final consonant (e.g., *cheab*) and those with a voiceless final consonant (e.g., *glope*). Although both types of near-words caused activation of the base words, there was more lexical activation after presentation of near-words like *glope* than after near-words like *cheab*. Near-words like *glope* caused as much lexical activation as real words (e.g., *globe*) did. As the level of lexical activation after presentation of a near-word was so high, such activation may not be easy to overcome, and it may greatly affect the comprehension of nonnative speech.

Importantly, an increase of lexical activation for nonnative listeners compared to native listeners was found both for near-words based on the vowel contrast and for those based on the consonant voicing contrast. Thus, it was not restricted to stimuli containing nonnative contrasts that were difficult to distinguish for the nonnative listeners.

Partially overlapping words

In Chapter 4, the perception of partially overlapping words was investigated. Word pairs were used which had similar onsets or fully overlapped, except for one phoneme contrast. Again, the items either differed in the /æ/-/ɛ/ contrast, which was difficult to distinguish for the Dutch listeners, or in the word-final consonant voicing contrasts, which the Dutch listeners distinguished as accurately as the English listeners did.

Dutch and English listeners' perception of words with partially overlapping onsets was assessed in the cross-modal priming task of Experiment 4.1. Pairs of trisyllabic words were used, the first parts of which were identical, except that one word contained an /æ/ and the other an /ɛ/, for example *daffodil* and *deficit*. Both for the Dutch listeners and for the English listeners, hearing the beginning of a word (e.g., *daffo*) facilitated the recognition of the corresponding word (*daffodil*). For the Dutch listeners, the beginning of that word also facilitated the recognition of the paired word. For these listeners, *daffo* activated both *daffodil* and *deficit*. For the English listeners on the other hand, there was no evidence that the beginning of one word activated the other word. Thus, there was more lexical activation of competitor words with partially overlapping onsets for the Dutch listeners than for the English listeners.

The perception of minimal pairs was investigated in the cross-modal priming task of Experiment 4.2. The word pairs differed only in the /æ/-/ɛ/ contrast (e.g., *flash - flesh*) or in the word-final consonant voicing contrast (e.g., *robe - rope*). Both for the English listeners and for the Dutch listeners, the presentation of one word (e.g., *flash*) facilitated the recognition of that same word (*flash*). For the English listeners, presentation of one word inhibited the recognition of the minimally different word (*flesh*). For the Dutch listeners on the other hand no facilitation nor inhibition of the minimally different word was found. This suggests that for the Dutch listeners, presentation of one word sometimes inhibited and sometimes facilitated recognition of the paired word. Thus, sometimes the target word inhibited its competitor, and sometimes both words remained active.

For the vowel items, the increased lexical activation for the nonnative listeners could be explained by the perceptual ambiguity of the /æ/-/ɛ/ contrast for these listeners. However, similar results were found for minimal pairs which differed in contrasts which were easy to distinguish for the nonnative listeners. Thus, increased lexical activation of partially overlapping words in nonnative listening was not restricted to perceptually ambiguous stimuli.

Conclusions

This dissertation investigated several aspects of phonetic and lexical processing in the perception of a second language. It provides new insights into the processing of nonnative speech. Further, the results from this study raise some interesting new questions.

First, this dissertation provides insight into the way listeners perceive nonnative phoneme contrasts that are similar to native contrasts, but in a position where they do not occur in the native language. Although the perception of nonnative phoneme contrasts has been studied widely (e.g., see Strange, 1995), not much is known about the role of the phonotactic constraints of the native language. Current models concerned with the perception of nonnative phonemes (Best, 1994; Flege, 1995) do not make specific predictions about the perception of nonnative but familiar contrasts in unfamiliar positions. The study described in Chapter 2 sheds some light on this novel issue. The results suggest that an unfamiliar position does not necessarily diminish perceptual accuracy.

Second, Chapter 2 discussed the necessity to use perceptual cues in a native-like manner for accurate perception of nonnative contrasts. The results showed that Dutch listeners did not use the duration of the preceding vowel as a cue to final obstruent voicing in a native-like manner. Nevertheless, they categorized the voicing contrasts with a native-like level of accuracy. Thus, a native-like use of perceptual cues is not always necessary for the accurate distinction of familiar contrasts in unfamiliar positions.

Third, this dissertation gives important insights in lexical processing during the comprehension of speech in a second language. Chapters 3 and 4 showed that the process of lexical activation is much more extensive in nonnative than in native listening. Several types of speech input caused an increase of lexical activation in nonnative as compared to native listening. For nonnative listeners, near-words caused more activation of the nearest word form, words with partially overlapping onsets remained ambiguous longer, and minimal pairs activated each other more than for native listeners. The study showed that spurious activation could be very strong. For example, the level of activation after presentation of a near-word was often as high as that after presentation of a real word.

Finally, the research described in Chapters 3 and 4 showed that increased lexical activation is not restricted to stimuli containing phoneme contrasts that are difficult to distinguish for nonnative listeners. It also occurred for items which contained no perceptually ambiguous phonemes. This increases the number of possible lexical competitors for nonnative listeners considerably. Thus, for Dutch listeners, presentation of the word *bad* may not only activate the lexical competitor *bed*, but it may activate *bat* and possibly even *bet* as well. Thus, the increased activation of lexical competitors may be highly pervasive in the comprehension of nonnative speech.

Future research could further examine the question which nonnative phoneme contrasts are likely to cause increased lexical activation for nonnative listeners. Increased lexical activation is not expected to be induced by all nonnative phonemes. In the present study, it was caused by familiar contrasts in an unfamiliar position, which were similar to contrasts in the native language, but occurred in a position where they are not contrastive in the native language. Thus, the phonotactic constraints of the native language may be an important factor in the occurrence of spurious lexical activation in a second language.

Further, the present results raise the question how nonnative contrasts can cause increased lexical activation when phoneme perception is uncompromised. One possible explanation is that nonnative listeners may disregard certain contrasts for lexical access, possibly induced by the phonotactics of the native language. Another possibility is that nonnative speakers may regularly mispronounce particular contrasts, which could affect the nonnative listeners' lexical representations, which could in turn lead to increased lexical activation. This issue deserves further research.

What does this research mean for the second-language learner? This dissertation contains both good and bad news for the second-language learner. On the positive side, phonetic processing does not have to be native-like for the nonnative listener to recognize sounds as accurately as native listeners do. On the negative side, even if phonetic processing is accurate, lexical processing may not be as efficient as it is in the native language, due to an increase in lexical activation in nonnative listening. Thus, this study adds another item to the long list of complications that make the comprehension of speech in a second language difficult.

References

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.

Best, C. T., McRoberts, G. W., & Sithole, N. M. (1988). Examination of perceptual reorganization for nonnative speech contrasts: Zulu click discrimination by English-speaking adults and infants. *Journal of Experimental Psychology: Human Perception and Performance, 14*, 345-360.

Booij, G. (1995). *The Phonology of Dutch*. Oxford: Oxford University Press.

Flege, J. E. (1995). Second language learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech Perception and Linguistic Experience: Issues in Cross-Language Research* (pp. 233-272). Baltimore: York Press.

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

Norris, D., McQueen, J. M., & Cutler, A. (2000). Merging information in speech recognition: Feedback is never necessary. *Behavioral and Brain Sciences, 23*, 299-325.

Pitt, M. A., & Samuel, A. G. (1993). An empirical and meta-analytic evaluation of the phoneme identification task. *Journal of Experimental Psychology: Human Perception and Performance, 19*, 699-725.

Strange, W. (Ed.). (1995). Speech Perception and Linguistic Experience: Issues in Cross-Language Research. Baltimore: York Press.

Samenvatting

Fonetische en lexicale processen in een tweede taal

Het verstaan van spraak in een tweede taal

Het verstaan van spraak in een tweede taal is veel moeilijker dan het verstaan van spraak in je moedertaal. Je moedertaal versta je meestal moeiteloos en luisteraars zijn zich dan ook zelden bewust van de complexe processen die hiervoor nodig zijn. Bij het luisteren naar een tweede taal worden mensen zich juist vaak pijnlijk bewust van deze complexiteit.

Wanneer je spraak in een tweede taal probeert te verstaan kan dit zoveel moeite kosten dat je denkvermogen er tijdelijk door afneemt. Dit bleek uit een onderzoek van Takano en Noda (1993). Deelnemers aan dit onderzoek voerden verschillende cognitieve taken uit, zoals rekenen en de weg zoeken uit een getekend doolhof. Tijdens deze taken luisterden ze naar spraak in hun moedertaal of naar spraak in een tweede taal die ze goed kenden. Wanneer de deelnemers luisterden naar spraak in een tweede taal maakten ze meer fouten in de cognitieve taken dan wanneer ze naar hun moedertaal luisterden. Het luisteren naar een tweede taal kostte blijkbaar zoveel moeite dat het de luisteraars hinderde bij het uitvoeren van andere cognitieve taken.

Het verstaan van een tweede taal is in allerlei opzichten moeilijk. Spraak in een tweede taal lijkt soms te snel om er losse woorden in te herkennen, sommige klanken zijn moeilijk uit elkaar te houden, er kunnen woorden en uitdrukkingen gebruikt worden die de luisteraar niet kent, het herkennen van woorden kost meer tijd dan in de moedertaal, en zelfs als de luisteraar alle afzonderlijke woorden heeft herkend is het nog niet altijd duidelijk wat de zin als geheel betekent.

Sommige problemen die een luisteraar tegen komt bij het verstaan van een tweede taal zijn gemakkelijk te begrijpen. Het is logisch dat je een kleiner vocabulaire hebt in een tweede taal dan in je moedertaal, vooral als je pas begonnen bent de taal te leren, en dat je daarom woorden kunt tegenkomen die je niet kent. Andere problemen zijn minder doorzichtig. Deze problemen kunnen worden verklaard aan de hand van de cognitieve processen waarop het verstaan van spraak is gebaseerd.

Het verdelen van spraak in afzonderlijke woorden

Bijvoorbeeld, waarom lijkt spraak in een tweede taal vaak sneller dan spraak in je moedertaal? Per definitie geldt dat vreemde talen niet altijd sneller kunnen zijn dan moedertalen, want wat een vreemde taal is voor de één is de moedertaal van een ander. Toch hebben mensen vaak de indruk dat er in een vreemde taal - denk bijvoorbeeld aan het Frans - heel snel gepraat wordt. De verklaring hiervoor is waarschijnlijk dat het luisteraars niet altijd lukt om spraak in een tweede taal te verdelen in afzonderlijke woorden, zodat ze de spraak horen als één lange, ononderbroken stroom. In geschreven taal zijn de woorden duidelijk gescheiden door een spatie. In spraak daarentegen is er geen vergelijkbare manier om woorden van elkaar te scheiden. Pauzes geven in spraak geen woordgrenzen aan zoals spaties op papier: pauzes komen vaak midden in woorden voor en ontbreken vaak tussen de woorden. Daarom moeten luisteraars andere informatie gebruiken om de grenzen van woorden in een gesproken zin te vinden. Luisteraars met verschillende moedertalen doen dit op verschillende manieren, en de manier om woordgrenzen te vinden in je moedertaal hoeft niet te werken in een andere taal.

Er zijn verschillende soorten strategieën die luisteraars kunnen gebruiken om lopende spraak in afzonderlijke woorden te verdelen. Eén groep strategieën is gebaseerd op de ritmische structuur van de taal. Deze zogenaamde metrische segmentatiestrategieën kunnen per taal verschillen. In het Nederlands en in het Engels beginnen de meeste woorden met een beklemtoonde lettergreep. Nederlandse en Engelse luisteraars gebruiken deze informatie (onbewust) om het begin van een woord in lopende spraak te vinden (Cutler & Norris, 1988; Vroomen, Van Zon, & De Gelder, 1996). Zij kunnen woorden die beginnen met een beklemtoonde lettergreep gemakkelijker vinden in lopende spraak dan woorden die beginnen met een onbeklemtoonde lettergreep. Franse en Spaanse luisteraars daarentegen gebruiken de grenzen van lettergrepen om woordgrenzen te vinden (Cutler, Mehler, Norris, & Seguí, 1986; Sebastián-Gallés, Dupoux, Seguí, & Mehler, 1992). In het Japans is de ritmische structuur niet gebaseerd op lettergrepen maar op de zogenaamde mora. De merknaam *Mazda* bestaat bijvoorbeeld uit drie mora's: ma-z-da. Japanse luisteraars gebruiken een metrische segmentatiestrategie die lijkt op die van de Franse en Spaanse luisteraars, maar ze gebruiken hierbij de grenzen van mora's in plaats van lettergrepen (Otake, Hatano, Cutler, & Mehler, 1993). Er zijn dus verschillende soorten metrische segmentatiestrategieën. Maar wanneer mensen naar een tweede taal luisteren gebruiken ze niet de strategie die bij die taal past, maar de strategie die ze in hun eigen moedertaal gebruiken (Cutler et al., 1986; Otake et al., 1993). Dit kan goed werken wanneer de metrische structuren van de moedertaal en de tweede taal overeenkomen, zoals voor het Nederlands en het Engels. Wanneer de metrische structuren van de twee talen verschillen is de strategie die werkt voor de moedertaal niet geschikt voor de tweede taal. In zulke gevallen is het moeilijk voor de luisteraar om de stroom van lopende spraak te verdelen in afzonderlijke woorden. Voor Nederlandse luisteraars kan het dus lijken alsof er in het Frans heel snel gesproken wordt omdat de Nederlandse metrische segmentatiestrategieën niet werken in het Frans.

Een ander type segmentatiestrategieën is gebaseerd op de zogenaamde fonotactische regels van een taal. Fonotactische regels bepalen welke spraakklanken en welke combinaties van klanken kunnen voorkomen en waar. Bijvoorbeeld, in het Engels kan een lettergreep beginnen met /sl/ (zoals in *sleep*), maar niet met /nl/. Wanneer Engelse luisteraars de klanken /nl/ achter elkaar horen zouden ze dus kunnen afleiden dat deze klanken deel uitmaken van verschillende lettergrepen (bijv. *unless*) en misschien ook van verschillende woorden (bijv. *on loan*). Er is inderdaad aangetoond dat Engelse luisteraars dit soort informatie gebruiken om woorden te vinden in lopende spraak (Weber, 2001). Woorden die met een /l/ begonnen waren voor hen gemakkelijker te vinden wanneer ze na een /n/ kwamen (/nl/) dan wanneer ze na een /s/ kwamen (/sl/). Duitse luisteraars die Engels als tweede taal kenden gebruikten ook Engelse fonotactische regels wanneer ze naar het Engels luisterden. Maar zij gebruikten daarnaast fonotactische regels die specifiek waren voor het Duits en die niet nuttig waren in het Engels (Weber, 2001). Deze inmenging van de fonotactische regels van de moedertaal kan het verdelen van lopende spraak in afzonderlijke woorden bemoeilijken bij het luisteren naar een tweede taal.

Het herkennen van spraakklanken

Een ander probleem is dat de spraakklanken van een tweede taal soms erg moeilijk uit elkaar te houden zijn. Een bekend voorbeeld hiervan zijn de klanken /r/ en /l/ voor Japanse luisteraars. Zij vinden het niet alleen moeilijk om deze klanken uit te spreken (Flege, Takagi, & Mann, 1995), maar ook om het verschil tussen de klanken te horen (Best & Strange, 1992). De verklaring hiervoor is dat er geen /r/ of /l/ is in het Japans, maar wel een klank die ergens tussen een /r/ en een /l/ inzit.

Interessant genoeg kunnen baby’s de klanken van andere talen vaak beter herkennen dan volwassenen. Aanvankelijk kunnen baby’s alle mogelijke spraakklanken heel goed uit elkaar houden. Dit geldt zelfs voor klanken van talen die ze nog nooit gehoord hebben. Maar wanneer ze de klanken van hun moedertaal beter leren kennen verliezen ze hun gevoeligheid voor de klanken van andere talen. Engelse baby’s konden bijvoorbeeld het verschil tussen twee medeklinkers uit het Hindi wel horen toen ze zes tot acht maanden oud waren, maar niet meer toen ze elf tot dertien maanden waren (Werker & Lalonde, 1988). Verder konden ze het verschil tussen twee Duitse klinkers horen toen ze vier maanden oud waren, maar niet meer toen ze zes maanden waren (Polka & Werker, 1994).

Dit betekent niet dat voor volwassen luisteraars alle klanken van een tweede taal moeilijk uit elkaar te houden zijn. Ook voor volwassenen zijn sommige van deze klanken nog steeds gemakkelijk te onderscheiden. Dit geldt bijvoorbeeld vaak voor klanken die sterk lijken op klanken uit de moedertaal (Best & Strange, 1992). Maar het tegendeel komt ook voor: klanken uit een tweede taal kunnen zo verschillen van de moedertaal dat ze niet eens als spraak herkend worden. Omdat deze klanken niet worden verward met de klanken van de moedertaal zijn ze gemakkelijk uit elkaar te houden. Zo zijn klikklanken uit het Zulu heel gemakkelijk van elkaar te onderscheiden voor Engelse luisteraars (Best, McRoberts, & Sithole, 1988). Daar tegenover staan spraakklanken die erg moeilijk te onderscheiden zijn. Het verschil tussen twee klanken uit een tweede taal is moeilijk te horen als er in de moedertaal maar één klank is die erop lijkt. Als één van de twee klanken uit de tweede taal meer op de klank uit de moedertaal lijkt dan de andere maakt dit het onderscheid weer een beetje gemakkelijker (Best & Strange, 1992). Wanneer beide klanken uit de tweede taal sterk lijken op de ene klank uit de moedertaal, zoals in het geval van /r/ en /l/ voor Japanse luisteraars, is het bijzonder moeilijk om de klanken uit elkaar te houden (Best & Strange, 1992).

Het herkennen van woorden

Ook het herkennen van woorden is moeilijker in een tweede taal dan in de moedertaal. Dit geldt natuurlijk voor woorden die de luisteraar niet goed kent, maar het geldt ook voor andere woorden. Om duidelijk te kunnen maken hoe dit komt is er eerst een korte uitleg nodig over de manier waarop woorden worden herkend.

Alle woorden die een luisteraar kent zijn opgeslagen in het brein, in het zogenaamde mentale lexicon. Als iemand een ander hoort praten worden er woorden in het mentale lexicon geactiveerd. Behalve de woorden die de spreker bedoelde (bijv. het woord *kapitein*), worden er ook woorden geactiveerd die hierop lijken (bijv. het woord *kapitaal*) (Zwitserlood, 1989). Zolang deze woorden in het mentale lexicon actief zijn 'strijden' ze met elkaar om herkend te worden. Deze strijd gaat door totdat er inderdaad één woord herkend wordt. Normaal is dat het woord dat de spreker ook daadwerkelijk gebruikte. Hoewel het omslachtig klinkt is dit een efficiënte manier om gesproken woorden te herkennen. Het heeft echter één nadeel. Hoe meer woorden er tegelijk actief zijn, hoe moeilijker het is om het juiste woord te herkennen (McQueen, Norris, & Cutler, 1994; Norris, McQueen, & Cutler, 1995).

Bij het verstaan van een tweede taal kunnen er niet alleen woorden uit die taal worden geactiveerd, maar ook woorden uit de moedertaal van de luisteraar. Dit is gebleken uit verschillende studies. Als Nederlandse luisteraars het Engelse woord *leaf* hoorden werd niet alleen dat woord geactiveerd in hun mentale lexicon, maar ook het Nederlandse woord *lief* (Schulpen, Dijkstra, Schriefers, & Hasper, 2003). Als ze het Engelse woord *desk* hoorden werd ook het Nederlandse woord *deksel* geactiveerd, dat op dezelfde manier begint (Weber & Cutler, 2004). Als luisteraars met Russisch als moedertaal het Engelse woord *marker* hoorden werd het Russische woord *marku* geactiveerd (Marian, Spivey, & Hirsch, 2003). Kortom, als mensen naar een tweede taal luisteren worden er ook woorden uit hun eerste taal geactiveerd. Omdat het moeilijker is om woorden te herkennen naarmate er meer andere woorden geactiveerd zijn, bemoeilijkt de activatie van woorden uit de moedertaal de herkenning van woorden in een tweede taal.

Het begrijpen van zinnen

Een laatste hindernis bij het verstaan van spraak is het begrijpen van de zin als geheel. Om de betekenis van een zin te kunnen begrijpen moet de luisteraar in elk geval sommige woorden hebben verstaan. Maar wat een zin betekent ligt voor een deel ook besloten in de zogenaamde prosodische informatie, zoals het zinsaccent.

De volgende zinnen (vertaald uit Akker & Cutler, 2003) verschillen alleen in het zinsaccent en hebben daardoor een verschillende betekenis: (1) *De toerist vloog NIET naar huis.* (2) *De toerist VLOOG niet naar huis.* Een studie liet zien dat Nederlandse en Engelse luisteraars op dezelfde manier gebruik maakten van prosodische informatie als ze naar hun moedertaal luisterden. Maar als Nederlandse luisteraars naar hun tweede taal Engels luisterden waren ze minder efficiënt in het gebruik van prosodische informatie (Akker & Cutler, 2003). Een minder efficiënt gebruik van prosodische informatie kan het begrijpen van zinnen in een tweede taal bemoeilijken.

Prosodische informatie kan ook worden gebruikt om te bepalen of een zin letterlijk of figuurlijk bedoeld is. Sommige zinnen, zoals *het ijs was gebroken* of *de kust is veilig* kunnen zowel letterlijk of figuurlijk worden geïnterpreteerd. Aan de manier waarop een zin wordt uitgesproken kunnen luisteraars in hun moedertaal horen of de spreker de letterlijke of de figuurlijke betekenis bedoelt (Vanlancker-Sidtis, 2003). In een tweede taal zijn luisteraars hier niet goed in. Zelfs luisteraars die hun tweede taal Engels heel goed beheersten bleken niet goed te zijn in het herkennen van letterlijke en figuurlijke betekenissen, en luisteraars die het Engels minder goed beheersten konden dit helemaal niet (Vanlancker-Sidtis, 2003). In een tweede taal zullen luisteraars daarom soms gebruik moeten maken van andere informatie, zoals de context, om letterlijk en figuurlijk bedoelde uitspraken als zodanig te herkennen.

Dit proefschrift

Onderzoeksvragen

Dit proefschrift gaat verder in op twee van de processen die moeilijk zijn bij het verstaan van spraak in een tweede taal, namelijk het herkennen van spraakklanken en het herkennen van woorden. De stappen die leiden tot het herkennen van spraakklanken worden fonetische processen genoemd en de stappen die leiden tot het herkennen van woorden worden lexicale processen genoemd. Vandaar de titel van dit proefschrift: 'Fonetische en lexicale processen in een tweede taal'.

Ten eerste is onderzocht in hoeverre Nederlandse luisteraars twee moeilijke Engelse spraakklanken herkenden. Het ging hierbij om de klinkers in de woorden *lamp* en *desk*, die fonetisch genoteerd worden als /æ/ en /ɛ/. De situatie is vergelijkbaar met die van de /r/ en de /l/ voor Japanse luisteraars. Het Nederlands beschikt niet over dezelfde /æ/ en /ɛ/ als het Engels, maar wel over een klank die daar tussenin zit, namelijk de klinker in het Nederlandse woord *pet*. De verwachting was dan ook dat deze Engelse klanken moeilijk te herkennen zouden zijn voor Nederlandse luisteraars. Daarom is onderzocht in hoeverre Nederlandse luisteraars de klanken uit elkaar konden houden en of ze dit minder goed konden dan luisteraars met Engels als moedertaal.

Ook is onderzocht in hoeverre Nederlandse luisteraars de Engelse klanken /z/ en /s/, /v/ en /f/, /b/ en /p/, en /d/ en /t/ konden herkennen. Deze Engelse klanken lijken sterk op Nederlandse klanken en zouden daarom gemakkelijk te herkennen kunnen zijn voor Nederlandse luisteraars. In het Nederlands worden /z/, /v/, /b/ en /d/ aan het eind van een woord echter uitgesproken als /s/, /f/, /p/ en /t/. Bijvoorbeeld, *honden* wordt uitgesproken met een /d/, maar *hond* met een /t/ aan het eind. In het Nederlands hoeven luisteraars daarom nooit onderscheid te maken tussen /d/ en /t/ aan het eind van een woord, en hetzelfde geldt voor de andere drie klankparen. In het Engels kunnen al deze klanken echter aan het eind van een woord voorkomen. Bijvoorbeeld, *robe* wordt uitgesproken met een /b/ en *rope* met een /p/. Voor Nederlandse luisteraars zou het moeilijk kunnen zijn om deze klanken aan het eind van een woord uit elkaar te houden in het Engels. Daarom is onderzocht in hoeverre Nederlandse luisteraars deze Engelse klanken konden herkennen aan het eind van een woord. Er is onderzocht of ze het moeilijker vonden om de klanken te herkennen wanneer ze aan het eind van een woord stonden dan wanneer ze aan het begin van een woord stonden, en of ze het moeilijker vonden om de klanken te herkennen dan luisteraars met Engels als moedertaal.

Ten tweede is de herkenning van woorden onderzocht. Zoals hierboven is uitgelegd is het moeilijker om gesproken woorden te herkennen naarmate er meer woorden geactiveerd zijn in het mentale lexicon. Wanneer mensen luisteren naar een tweede taal worden er ook woorden uit hun eerste taal geactiveerd, wat de herkenning van de gesproken woorden bemoeilijkt. In dit proefschrift is onderzocht of er voor mensen die luisteren naar een tweede taal ook meer woorden uit die tweede taal geactiveerd worden dan voor mensen die naar hun moedertaal luisteren. Als een Nederlandse en een Engelse luisteraar dezelfde Engelse spraak horen zouden er dus meer Engelse woorden geactiveerd kunnen worden voor de Nederlandse luisteraar dan voor de Engelse luisteraar. Dit zou de herkenning van woorden in een tweede taal kunnen bemoeilijken. Deze extra activatie van Engelse woorden voor Nederlandse luisteraars zou in verschillende situaties kunnen optreden:

Als een Nederlandse luisteraar iemand *daf* hoorde zeggen zou hij gemakkelijk kunnen denken dat hij het woord *deaf* had gehoord. Op dezelfde wijze zouden *lemp* en *glite* kunnen worden verstaan als *lamp* en *glide*. Items als *daf, lemp* en *glite*, die geen bestaande woorden vormen maar wel erg lijken op bestaande woorden, worden in dit proefschrift *near-words*, ofwel 'bijna-woorden', genoemd. Uiteraard gebruiken Engelse sprekers deze bijna-woorden nooit zomaar, omdat het geen bestaande woorden zijn. Maar Engelse sprekers kunnen wel het woord *DAFfodil* gebruiken, dat met het bijna-woord *daf* begint. Ze kunnen het hebben over het *eviL EMPire*, waarin het bijna-woord *lemp* is ingebed, of over een *biG LIGHT*, dat het bijna-woord *glite* bevat. Wanneer luisteraars een bijna-woord zoals *daf* horen kan een woord als *deaf* geactiveerd worden in het mentale lexicon. Dit zou vaker kunnen gebeuren voor Nederlandse luisteraars dan voor Engelse luisteraars, bijvoorbeeld doordat de Nederlandse luisteraars het verschil tussen het bijna-woord en het woord niet goed kunnen horen (zoals bij *daf* en *deaf*), of doordat ze geen aandacht besteden aan het verschil (zoals bij *glite* en *glide*). Daarom is onderzocht of Engelse bijna-woorden meer activatie van woorden in het mentale lexicon veroorzaakten voor Nederlandse luisteraars dan voor Engelse luisteraars.

Verder is de herkenning onderzocht van woorden die bijna hetzelfde zijn of die op bijna dezelfde manier beginnen. Bijvoorbeeld, de eerste delen van de woorden *daffodil* en *deficit* klinken bijna hetzelfde. Wanneer een luisteraar het eerste deel van *daffodil* hoort zou dit kunnen leiden tot de activatie van *deficit* in het mentale lexicon. Dit zou meer kunnen gebeuren voor Nederlandse luisteraars dan voor Engelse luisteraars, aangezien Nederlandse luisteraars het verschil misschien minder goed kunnen horen. De woorden *flash* en *flesh*, of *robe* en *rope* lijken nog meer op elkaar, en het is dan ook mogelijk dat Nederlandse luisteraars deze woorden helemaal niet uit elkaar houden. Hetzelfde geldt voor de woorden *cattle* ('vee') en *kettle* ('ketel'), die staan afgebeeld op de omslag van dit proefschrift. Er is onderzocht in hoeverre het horen van het eerste deel van een woord als *daffodil* of van een heel woord als *flash* activatie veroorzaakt van het woord dat erop lijkt, en of dit meer is voor Nederlandse luisteraars dan voor Engelse luisteraars. Als dit het geval is, zouden er vaak meer Engelse woorden geactiveerd worden voor Nederlandse luisteraars dan voor Engelse luisteraars bij het horen van dezelfde Engelse spraak. Dit zou de herkenning van woorden in een tweede taal bemoeilijken.

Resultaten en conclusies

De herkenning van de klanken /æ/ en /ɛ/ is bestudeerd in zes experimenten die staan beschreven in de hoofdstukken 2, 3 en 4. In alle experimenten bleken Nederlandse luisteraars de klanken minder goed te herkennen dan Engelse luisteraars. De Nederlandse luisteraars herkenden de klanken wel boven kansniveau, wat wil zeggen dat hun resultaten beter waren dan wanneer ze alleen maar hadden gegokt. In Experiment 2.1 benoemden de Nederlandse luisteraars de klanken correct in maar liefst 95 % van de gevallen (tegenover 99 % voor de

Engelse luisteraars), in de andere experimenten slechts in gemiddeld 59 % van de gevallen (tegenover 84 % voor de Engelse luisteraars).

De herkenning van de klanken /z/-/s/, /v/-/f/, /b/-/p/ en /d/-/t/ is bestudeerd in vijf experimenten, beschreven in de hoofdstukken 2, 3 en 4. De Nederlandse luisteraards bleken de klanken aan het eind van een woord net zo goed te herkennen als die aan het begin van een woord, en net zo goed als de Engelse luisteraars dit deden. Uit Experiment 2.2 bleek wel dat de Nederlanders zich bij het herkennen van de klanken niet op dezelfde kenmerken van het spraaksignaal baseerden als de Engelse luisteraars.

De experimenten in Hoofdstuk 3 onderzochten of het horen van Engelse bijna-woorden (bijv. *lemp* of *glite*) meer activatie veroorzaakte van woorden in het mentale lexicon (bijv. *lamp* of *glide*) voor Nederlandse luisteraars dan voor Engelse luisteraars. Dit bleek inderdaad het geval te zijn. Voor de Nederlandse luisteraars was er vaak net zoveel activatie van een woord in het mentale lexicon na het horen van een bijna-woord als na het horen van een echt woord (Experiment 3.2 en 3.3). Nederlandse luisteraars dachten na het horen van een bijna-woord vaker dat ze een echt woord hadden gehoord dan Engelse luisteraars (Experiment 3.1). Dit gold zowel voor bijna-woorden die van echte woorden verschilden in klanken die voor de Nederlandse luisteraars moeilijk te onderscheiden waren (bijv. *lemp* - *lamp*) als voor bijna-woorden die van echte woorden verschilden in klanken die gemakkelijk te onderscheiden waren (bijv. *glite* - *glide*).

De experimenten in Hoofdstuk 4 onderzochten de herkenning van woordparen die bijna hetzelfde waren of die op bijna dezelfde manier begonnen. In Experiment 4.1 werden woordparen gebruikt die op bijna dezelfde manier begonnen, zoals *daffodil* en *deficit*. Als Engelse luisteraars het begin van het woord *daffodil* hoorden was daarna het woord *daffodil* in het mentale lexicon actief, maar niet het woord *deficit*. Voor Nederlandse luisteraars waren echter beide woorden actief. In Experiment 4.2 werden woordparen gebruikt zoals *flash* en *flesh* of *robe* en *rope*. Voor de Engelse luisteraars bleef na het horen van een woord als *flash* alleen dat woord actief in het mentale lexicon en het andere woord (*flesh*) niet. Voor de Nederlandse luisteraars bleven in sommige gevallen beide woorden actief. Ook hierbij was er dus voor Nederlandse luisteraars meer activatie van Engelse woorden in het mentale lexicon dan voor Engelse luisteraars.

Er zijn in dit proefschrift een aantal verschillen gevonden tussen de manier waarop mensen klanken en woorden herkennen in hun moedertaal en in een tweede taal. Wat betekenen deze bevindingen voor de luisteraar? Voor de luisteraar bevat dit proefschrift zowel goed als slecht nieuws. Het goede nieuws is dat luisteraars klanken uit een tweede taal heel goed uit elkaar kunnen houden, zelfs als deze voorkomen op plaatsen waar ze in de moedertaal niet voorkomen (zoals /d/ aan het eind van een woord). Het slechte nieuws is dat het herkennen van woorden in een tweede taal minder efficiënt kan verlopen dan in de moedertaal, doordat er meer woorden van de tweede taal actief zijn in het mentale lexicon. Het onderzoek in dit proefschrift voegt dan ook een nieuw punt toe aan de lange lijst met complicaties die het verstaan van spraak in een tweede taal bemoeilijken.

Referenties

Akker, E., & Cutler, A. (2003). Prosodic cues to semantic structure in native and nonnative listening. *Bilingualism: Language and Cognition, 6*, 81-96.

Best, C. T., McRoberts, G. W., & Sithole, N. M. (1988). Examination of perceptual reorganization for nonnative speech contrasts: Zulu click discrimination by English-speaking adults and infants. *Journal of Experimental Psychology: Human Perception and Performance, 14*, 345-360.

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

Cutler, A., Mehler, J., Norris, D., & Seguí, J. (1986). The syllable's differing role in the segmentation of French and English. *Journal of Memory and Language, 25*, 385-400.

Cutler, A., & Norris, D. (1988). The role of strong syllables in segmentation for lexical access. *Journal of Experimental Psychology: Human Perception and Performance, 14*, 113-121.

Flege, J. E., Takagi, N., & Mann, V. (1995). Japanese adults can learn to produce English /ɹ/ and /l/ accurately. *Language and Speech, 38*, 25-55.

Marian, V., Spivey, M., & Hirsch, J. (2003). Shared and separate systems in bilingual language processing: Converging evidence from eyetracking and brain imaging. *Brain and Language, 86*, 70-82.

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.

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.

Otake, T., Hatano, G., Cutler, A., & Mehler, J. (1993). Mora or syllable? Speech segmentation in Japanese. *Journal of Memory and Language, 32*, 258-278.

Polka, L., & Werker, J. F. (1994). Developmental changes in perception of nonnative vowel contrasts. *Journal of Experimental Psychology: Human Perception and Performance, 20*, 421-435.

Schulpen, B., Dijkstra, T., Schriefers, H. J., & Hasper, M. (2003). Recognition of interlingual homophones in bilingual auditory word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 29*, 1155-1178.

Sebastián-Gallés, N., Dupoux, E., Seguí, J., & Mehler, J. (1992). Contrasting syllabic effects in Catalan and Spanish. *Journal of Memory and Language, 31*, 18-32.

Takano, Y., & Noda, A. (1993). A temporary decline of thinking ability during foreign language processing. *Journal of Cross-Cultural Psychology, 24*, 445-462.

Vanlancker-Sidtis, D. (2003). Auditory recognition of idioms by native and nonnative speakers of English: It takes one to know one. *Applied Psycholinguistics, 24*, 45-57.

Vroomen, J., Van Zon, M., & De Gelder, B. (1996). Cues to speech segmentation: Evidence from juncture misperception and word spotting. *Memory and Cognition, 24*, 744-755.

Weber, A. (2001). *Language-specific listening: The case of phonetic sequences.* Unpublished doctoral dissertation, Nijmegen University, The Netherlands.

Weber, A., & Cutler, A. (2004). Lexical competition in non-native spoken-word recognition. *Journal of Memory and Language, 50*, 1-25.

Werker, J. F., & Lalonde, C. E. (1988). Cross-language speech perception: Initial capabilities and developmental change. *Developmental Psychology, 24*, 672-683.

Curriculum vitae

Levensloop

Mirjam Broersma werd in 1976 in Groningen geboren. Ze doorliep de CBS De Windroos in Zuidhorn en het Praedinius Gymnasium te Groningen. In 1994 vertrok ze als au pair naar Parijs. Na terugkomst in Nederland begon ze in 1995 haar studie aan de Katholieke Universiteit Nijmegen. In 1996 behaalde ze haar propedeuse Arabisch (of, volgens het getuigschrift: 'Arabische, Nieuwperzische en Turkse Talen en Culturen'), waarna ze verder ging met een bovenbouwstudie Toegepaste Taalwetenschap. In 1999 studeerde ze in het kader van een uitwisselingsprogramma enkele maanden aan de University of South Carolina in Columbia (VS). In 2000 rondde ze haar studie *cum laude* af en werd haar scriptie beloond met de Anéla Scriptieprijs. Datzelfde jaar begon ze op het Max Planck Instituut voor Psycholinguïstiek te Nijmegen aan haar promotieonderzoek. In het kader van dit onderzoek en ondersteund door een Marie Curie Fellowship was ze in 2003 enige maanden verbonden aan het Laboratory of Experimental Psychology aan de University of Sussex in Brighton (GB). Momenteel is ze als postdoc werkzaam aan het Max Planck Instituut voor Psycholinguïstiek.

Publicaties / publications

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

Broersma, M. (2002). *Phoneme processing and lexical competition in the comprehension of non-native speech.* Poster presented at the Tutorials in Behavioral and Brain Sciences 2002, Kochel am See, Germany.

Broersma, M. (2005). Perception of familiar contrasts in unfamiliar positions. *Journal of the Acoustical Society of America, 117*, 3890-3901. *Chapter 2 of this dissertation.*

Broersma, M. (submitted). Competition increase in nonnative listening. *Language and Speech. Chapter 4 of this dissertation*.

Broersma, M. (submitted). Spurious lexical activation in nonnative listening. *Cognition. Chapter 3 of this dissertation*.

Broersma, M., & Cutler, A. (2001). *Comprehension of non-native speech: Inaccurate phoneme processing and activation of lexical competitors.* Poster presented at the Tutorials in Behavioral and Brain Sciences 2001, Nijmegen, The Netherlands.

Broersma, M., & De Bot, K. (2001). De triggertheorie voor codewisseling: De oorspronkelijke en een aangepaste versie. *Toegepaste Taalwetenschap in Artikelen, 65*(1), 41-54.

Broersma, M., & De Bot, K. (2006). Triggered codeswitching: A corpus-based evaluation of the original triggering hypothesis and a new alternative. *Bilingualism: Language and Cognition, 9*, in press.

Broersma, M., & Kolkman, K. M. (2004). *Lexical representation of non-native phonemes*. Proceedings of the 8th International Conference on Spoken Language Processing (pp. 1241-1244), Sunjin, Korea (CD-ROM).

Cutler, A., & Broersma, M. (2005). Phonetic precision in listening. In W. Hardcastle & J. Beck (Eds.), *A Figure of Speech* (pp. 63-91). Mahwah, NJ: Erlbaum.

MPI SERIES IN PSYCHOLINGUISTICS

1. The electrophysiology of speaking: Investigations on the time course of semantic, syntactic, and phonological processing. *Miranda van Turennout*
2. The role of the syllable in speech production: Evidence from lexical statistics, metalinguistics, masked priming, and electromagnetic midsaggital articulography. *Niels O. Schiller*
3. Lexical access in the production of ellipsis and pronouns. *Bernadette M. Schmitt*
4. The open-/closed-class distinction in spoken-word recognition. *Alette Haveman*
5. The acquisition of phonetic categories in young infants: A self-organising artificial neural network approach. *Kay Behnke*
6. Gesture and speech production. *Jan-Peter de Ruiter*
7. Comparative intonational phonology: English and German. *Esther Grabe*
8. Finiteness in adult and child German. *Ingeborg Lasser*
9. Language input for word discovery. *Joost van de Weijer*
10. Inherent complement verbs revisited: Towards an understanding of argument structure in Ewe. *James Essegbey*
11. Producing past and plural inflections. *Dirk Janssen*
12. Valence and transitivity in Saliba: An Oceanic language of Papua New Guinea. *Anna Margetts*
13. From speech to words. *Arie van der Lugt*
14. Simple and complex verbs in Jaminjung: A study of event categorisation in an Australian language. *Eva Schultze-Berndt*
15. Interpreting indefinites: An experimental study of children's language comprehension. *Irene Krämer*
16. Language-specific listening: The case of phonetic sequences. *Andrea Weber*
17. Moving eyes and naming objects. *Femke van der Meulen*
18. Analogy in morphology: The selection of linking elements in Dutch compounds. *Andrea Krott*

19. Morphology in speech comprehension. *Kerstin Mauth*
20. Morphological families in the mental lexicon. *Nivja H. de Jong*
21. Fixed expressions and the production of idioms. *Simone A. Sprenger*
22. The grammatical coding of postural semantics in Goemai (a West Chadic language of Nigeria). *Birgit Hellwig*
23. Paradigmatic structures in morphological processing: Computational and cross-linguistic experimental studies. *Fermín Moscoso del Prado Martín*
24. Contextual influences on spoken-word processing: An electrophysiological approach. *Daniëlle van den Brink*
25. Perceptual relevance of prevoicing in Dutch. *Petra M. van Alphen*
26. Syllables in speech production: Effects of syllable preparation and syllable frequency. *Joana Cholin*
27. Producing complex spoken numerals for time and space. *Marjolein Meeuwissen*
28. Morphology in auditory lexical processing: Sensitivity to fine phonetic detail and insensitivity to suffix reduction. *Rachèl J. J. K. Kemps*
29. At the same time...: The expression of simultaneity in learner varieties. *Barbara Schmiedtová*
30. A grammar of Jalonke argument structure. *Friederike Lüpke*
31. Agrammatic comprehension: An electrophysiological approach. *Marlies Wassenaar*
32. The structure and use of shape-based noun classes in Miraña (North West Amazon). *Frank Seifart*
33. Prosodically-conditioned detail in the recognition of spoken words. *Anne Pier Salverda*