COMPREHENSION OF NON-NATIVE SPEECH: INACCURATE PHONEME PROCESSING AND ACTIVATION OF LEXICAL COMPETITORS

Mirjam Broersma

Max Planck Institute for Psycholinguistics
Nijmegen, The Netherlands
mirjam.broersma@mpi.nl

ABSTRACT

Native speakers of Dutch with English as a second language and native speakers of English participated in an English lexical decision experiment. Phonemes in real words were replaced by others from which they are hard to distinguish for Dutch listeners. Non-native listeners judged the resulting near-words more often as a word than native listeners. This not only happened when the phonemes that were exchanged did not exist as separate phonemes in the native language Dutch, but also when phoneme pairs that do exist in Dutch were used in word-final position, where they are not distinctive in Dutch. In an English bimodal priming experiment with similar groups of participants, word pairs were used which differed in one phoneme. These phonemes were hard to distinguish for the non-native listeners. Whereas in native listening both words inhibited each other, in non-native listening presentation of one word led to unresolved competition between both words. The results suggest that inaccurate phoneme processing by non-native listeners leads to the activation of spurious lexical competitors.

1. INTRODUCTION

The vocabulary people have at their disposal in a second language is smaller than that in their native language. One could therefore assume that a smaller set of words is activated when listening to a second language. This would make lexical competition and selection easier and faster. There are two things that make this an unlikely assumption.

In the first place, evidence suggests that the native language cannot be entirely deactivated while listening to a second language. A series of visual lexical decision experiments, reviewed in [1], demonstrates that words from the native language are activated and compete for selection with words from the second language. An eyetracking experiment [2] showed that the lexicons of both languages of the bilingual subjects were active while they were listening to monolingual spoken stimuli. In another eyetracking experiment [1], bilingual subjects' eye movements to a set of pictures were monitored, while they were listening to spoken stimuli in their non-native language. Pictures with a native language name that was phonologically related to the non-native stimulus were fixated more than pictures with phonologically unrelated names. This outcome suggests that the names of the pictures in the native language were active, although subjects were listening to the non-native language. It seems therefore that the competitor set during the comprehension of non-native speech does not only consist of words from the relatively small vocabulary of this language, but also of words from the native language.

In the second place, people often have trouble distinguishing phonemes in a second language. Inaccuracies at the phoneme level may lead to the activation of spurious lexical competitors, which native listeners would not activate following the same speech input. These spurious competitors would enlarge the competitor set for non-native listeners, compared to that of native listeners, making the comprehension process harder.

Much research reveals that adult listeners show poor discrimination and identification of phones which are not linguistically distinctive in their native language (see [3] for a review). According to the perceptual assimilation model [4], listeners assimilate non-native phones to native phoneme categories, assigning them to the category to which they are phonetically most similar. The extent to which non-native contrasts can be distinguished is predicted on the basis of their occurrence and phoneme status in the native language.

Two experiments were carried out in order to test the relation between inaccurate phoneme processing and the activation of lexical competitors, in the case of native speakers of Dutch listening to their second language English.

The English vowels [ɛ] and [æ] should be highly confusable for Dutch listeners, since Dutch has the phoneme [ɛ] but not [æ]. Furthermore, we hypothesized that not only the phoneme status of a phone, but also the salience of features in certain positions in the word might determine the ease with which two phones can be distinguished. Although voiced and voiceless obstruents are separate phonemes in Dutch, voicing is not distinctive in word-final position, since word-final obstruents are always voiceless in Dutch. Therefore, the voiced and voiceless phoneme pairs [d] and [t], [v] and [f], [z] and [s], and [b] and [p] may be hard to distinguish in word-final position for Dutch listeners. In English on the other hand, voicing of obstruents is distinctive in all positions.

2. EXPERIMENT 1

In a lexical decision experiment, Dutch and English listeners were presented with so called 'near-words': non-words which were identical to existing English words, except that one phoneme was replaced with another with which it was expected to be highly confusable for Dutch listeners. Those confusable pairs were the vowels [ɛ] and [æ], and voiced and voiceless obstruents in word-final position.

2.1. Subjects

Twenty-four native speakers of Dutch from the Max Planck Institute participant pool and 24 native speakers of British English, recruited at the University of Birmingham, took part in this experiment. The Dutch participants had a high level of proficiency in their second language English. The English participants did not know any Dutch. None of the participants reported any hearing impairment. They received a small payment for their participation.

2.2. Materials

Sixty-four monosyllabic English words were selected as experimental items. Half of them contained an obstruent in word-final position (4 for every phoneme), the other half contained one of the target vowels [ɛ] and [æ] (16 per vowel). Further criteria for selection were that the word did not sound like an existing Dutch word, and that replacement of the target phoneme with its confusable counterpart did not result in an existing English or Dutch word. Near-words were formed by replacing the target phoneme with its confusable counterpart (e.g. 'desk', *dask*, 'globe', *glope*). The target words were divided into two lists, which were balanced for lemma frequency, with equal numbers of each target sound in each list. Each subject heard the 32 words from one list in their real-word form and the 32 words from the other list in near-word form. Sixty-eight monosyllabic English words, balanced for frequency with the target items, and 68 non-words, formed by replacing one phoneme in a real word, were selected as filler items.

The materials were recorded by a male native speaker of British English in a sound-attenuated booth onto digital audiotape and digitized at a sampling rate of 16 kHz.

2.3. Procedure

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. Items were presented in a random order, which was different for every participant. The randomization was such that minimally three other items appeared between two target words and between two near-words. The materials were presented binaurally over headphones. NESU software controlled the experiment. Participants were tested one at a time in a quiet room.

Participants 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 if they thought the presented item was an English word and a red response button if they thought it was not. The green response button was pressed with the dominant hand. Participants were asked to respond both as fast and as accurately as possible. The experiment started with 10 practice trials, after which participants had the opportunity to ask questions. The computer stored the reactions and reaction times. After each button press, the next trial was initialized. Presentation of an item started 500 ms after the response was given.

2.4. Results and discussion

Responses faster than 100 ms or slower than 2000 ms after the stimulus offset were discarded from the analysis. One experimental pair had to be excluded because one word inadvertently also appeared as a filler word.

Analyses of variance were carried out, separately across participants (F1) and across items (F2).

Figure 1 shows the percentages of near-words misjudged as words, and the percentages of their real-word counterparts judged as such. Both Dutch and English participants produced more yes-responses to real words than to near-words [Dutch: $F1\ (1,23) = 98.98$, $p < .000$, $F2\ (1, 62) = 82.08$, $p < .000$; English: $F1\ (1,23) = 2050.21$, $p < .000$, $F2\ (1,62) = 343.44$, $p < .000$]. The Dutch participants however showed a significantly higher percentage of yes-responses to near-words, relative to the real words, than the English [$F1(1,46) = 154.67$, $p < .000$, $F2\ (1, 124) = 68.97$, $p < .000$].

When the targets with vowel manipulation and those with consonant manipulation were analyzed separately, both sets showed this difference between the two language groups: Dutch participants had a significantly larger number of misjudgments of near-words, relative to the number of words judged as such, than English participants, both for vowel targets [$F1\ (1,46) = 74.35$, $p < .000$, $F2\ (1,62) = 29.01$, $p < .000$] and for consonant targets [$F1\ (1,46) = 78.43$, $p < .000$, $F2\ (1,60) = 62.71$, $p < .000$]. Both Dutch and English participants performed better on targets with consonant manipulation than on targets with vowel manipulation [Dutch: $F1\ (1,46) = 11.38$, $p < .002$, $F2\ (1,61) = 10.16$, $p < .002$; English: $F1\ (1,46) = 72.78$, $p < .000$, $F2\ (1, 61) = 20.44$, $p < .000$].

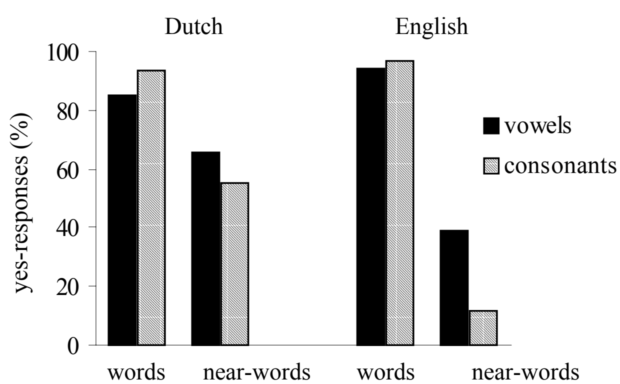

Figure 1: Percentages of words and near-words with vowel and consonant targets judged as words.

3. EXPERIMENT 2

In a bimodal priming experiment, Dutch and English listeners made lexicality judgements on English words which were preceded by unrelated, identical, or phonologically related prime words. Identical primes were expected to lead to a faster recognition of the target than unrelated primes. Related primes are normally predicted to inhibit the target word and lead to longer RTs. In this experiment, related primes only deviated from the target words in one phoneme, involving a contrast which was presumably hard to distinguish for Dutch listeners.

The same confusable phoneme pairs were used as in Experiment 1. Whereas the English participants were predicted to have longer RTs in the related condition compared to the unrelated control condition due to inhibition, for the Dutch listeners less inhibition, or even priming was predicted.

3.1. Participants

Seventy-two native speakers of Dutch from the Max Planck Institute participant pool and 72 native speakers of British English, recruited at the University of Birmingham, took part in this experiment. The Dutch participants had a high level of proficiency in their second language English. The English participants did not know any Dutch. None of them had participated in Experiment 1. They received a small payment for their participation.

3.2. Materials

Twenty-one pairs of English words were selected. Pairs were identical except for one phoneme. In 6 pairs, one word contained the vowel [ɛ] while the other contained [æ], in 10 pairs the target phonemes were [d] and [t], in 3 pairs [z] and [s], and in 2 pairs [b] and [p] (see Table 1 for examples). All target consonants occurred in word-final position. No pairs could be found with the target phonemes [v] and [f]. Four of the vowel targets were disyllabic, all other targets were monosyllabic.

For each pair, a semantically and phonetically unrelated word was chosen to function as prime in the unrelated condition. Each unrelated prime had the same number of syllables as its target pair and a lemma frequency matched to the frequencies of the two target words.

Target	Related prime	Unrelated prime
flesh	flash	spite
bride	bright	shave
phase	face	home
robe	rope	youth

Table 1: Examples of stimulus sets.

The sets of words were divided into three lists, balancing lemma frequencies, and such that the different target phonemes were distributed as equally as possible. Within each list, there were two groups, such that half of the words with a lemma frequency higher than that of their counterparts were grouped with half of the words with a frequency lower than that of their counterparts. Words from one group would be presented as targets to half of the subjects, words from the other group to the other half.

Twenty-one pairs of semantically and phonetically unrelated filler words were chosen for the unrelated condition, 21 pairs of related filler words (phonologically identical except for one phoneme) for the related condition, and 21 single filler words for the identical condition. Forty-two words were chosen to function as a related prime for a non-word which deviated from this word by one phoneme. Forty-two words were chosen as a prime for an unrelated non-word.

None of the items were phonetically similar to any Dutch words, and none of the words that were used as visual targets were orthographically similar to Dutch words.

The materials were recorded by the same speaker and in the same fashion as for Experiment 1.

3.3. Procedure

All participants were presented with one member of each experimental pair in one of three conditions; the prime was either the same word (identical condition: 'flesh'-'flesh'), the other member of the pair (related condition: 'flash'-'flesh') or an unrelated word (unrelated condition: 'spite'-'flesh'). All participants were presented with all filler items. Items were presented in a random order, which was different for every participant.

Prime words were presented binaurally over headphones, followed by the visual presentation of a target word on a computer screen directly after offset. NESU software controlled the experiment. Participants were tested one at a time in a quiet room.

Participants received written instructions in their native language, in which they were asked to decide if the visually presented word was an existing English word, and to indicate their response through button press. The experiment started with 12 practice trials. Otherwise the procedure was as in Experiment 1.

3.4. Results and discussion

RTs diverging by more than 3 standard deviations from the subject or item means were discarded from the analysis.

Analyses of variance were carried out across participants and across items.

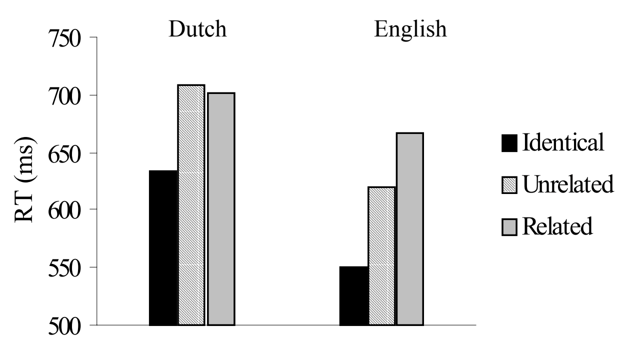

Figure 2: Reaction times for words preceded by identical, unrelated, and related primes.

Figure 2 shows RTs for experimental items preceded by identical, unrelated and related primes. Both Dutch and English participants responded faster in the identical condition than in the unrelated control condition [Dutch: F1 (1,71) = 27.23, p < .000, F2 (1,41) = 24.69, p < .000; English: F1 (1,71) = 60.94, p < .000, F2 (1,41) = 51.34, p < .000]. No significant difference between the language groups was found [F1 (1,142) = 0.24, p < .629, F2 (1,82) = 0.09, p < .761].

There was a significant difference however between the language groups for RTs in the related condition relative to the unrelated control condition [F1 (1,142) = 7.46, p < .007, F2

(1,82) = 6.63, p < .012]. English participants had longer RTs in the related condition than in the unrelated condition [F1 (1,71) = 20.51, p < .000, F2 (1,41) = 18.63, p < .000]. For the Dutch participants on the other hand, RTs in both conditions did not differ significantly [F1 (1,71) = 0.13, p < .718, F2 (1,41) = 0.04, p < .835]. RTs for Dutch participants in the related condition were longer than those in the identical condition [F1 (1,71) = 22.43, p < .000, F2 (1,41) = 31.49, p < .000]. Whereas related primes led to inhibition of the target for the English participants, no inhibition was found for the Dutch participants. We suggest that for them the competition between prime and target words remained unresolved. Neither one of the lexical competitors obtained sufficient activation to inhibit the other. Since RTs in the related condition were longer than those in the identical condition, primes must have activated identical targets more strongly than related targets. This in turn must mean that the non-native listeners succeeded in differentiating between the English phonemes, but not as clearly as the native listeners did. We suggest that the phonemic categories for non-native listeners may not be as distinct as those of the native listeners; indeed, they may even overlap, which would allow for (weaker) support to be given to the minimally mismatching words. Thus 'flesh' would activate 'flesh' more than 'flash' for native and non-native listeners alike; but while for native listeners 'flesh' would decisively mismatch 'flash', for non-native listeners 'flash' would be partially matched by 'flesh'. The mismatch of 'flash' would then not be large enough to lead to inhibition.

4. CONCLUSIONS

In Experiment 1, non-native listeners judged a non-word more often as a word than native listeners, when it involved a contrast which was hard to distinguish for the non-native listeners. Phoneme confusion leading to this misjudgment of near-words was not only caused by the lack of a phoneme distinction in the native language, but also by the lack of distinctiveness of a phoneme pair in word-final position in the native language.

In Experiment 2, words which were identical except for one phoneme inhibited each other in native listening, whereas presentation of one such word led to the activation of and unresolved competition between both words in non-native listening.

The results support the proposal that inaccurate phoneme processing by non-native listeners leads to the activation of spurious lexical competitors.

5. ACKNOWLEDGEMENTS

I thank Anne Cutler for suggesting these experiments and for her advice. Thanks also to Antje Meyer and the Behavioral Brain Science Center, University of Birmingham, for the use of their facilities.

6. REFERENCES

[1] Weber, A., Language-specific listening: The case of phonetic sequences, Nijmegen University, Nijmegen, 2001.

[2] Spivey, M. J., & Marian, V., "Cross talk between native and second languages: Partial activation of an irrelevant lexicon" *Psychological Science*, 10, 281-284, 1999.

[3] Werker, J. F., "Cross-language speech perception: Development change does not involve loss", in J. C. Goodman and H. C. Nusbaum (Eds.), The development of speech perception: The transition from speech sounds to spoken words (pp. 93-120). MIT, Cambridge, MA, 1994.

[4] Best, C. T., McRoberts, G. W., and Sithole, N. M., "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, 1988.