ON THE PERCEPTUAL RELEVANCE OF DEGEMINATION IN DUTCH

S.M.M. te Riele, M. Loef and O. van Herwijnen
Utrecht Institute of Linguistics (former OTS)
University of Utrecht, Trans 10, 3512 JK Utrecht, The Netherlands
Tel. +31 30 253 61 10, FAX +31 30 253 60 00, E-mail: Saskia.teRiele@let.ruu.nl

ABSTRACT

The aim of this study is to determine whether the phonological process of degemination, in which one of two adjacent and identical consonants is deleted, is perceptually complete when it applies over word boundaries. Measurements on the duration of the boundary consonant have shown durational differences between two-word phrases with underlying single and double consonants, even at fast speech rates. Results of a pseudo-gating experiment using a binary forced choice task show that correct segmentation of two-word phrases with underlying single or double consonants, spoken at a fast speech rate, does not exceed chance level. We conclude, therefore, that degemination actually occurs in Dutch and that this process is perceptually complete. Implications for word recognition will also be discussed.

1. INTRODUCTION

When words are combined into utterances, the phonetic word form may be modified, as a result of sandhi phenomena. Thus, a difference in sound structure between a word spoken in isolation and in combination with other words may occur. Examples of these kinds of modifications are assimilation, deletion and degemination. Consonants at word boundaries are especially sensitive to these processes.

In the present research we will concentrate on the process of degemination. The degemination rule in Dutch says that one of two adjacent and identical consonants in the underlying form is deleted in the phonetic surface representation. An example is the two-word phrase *zeef fijn* 'sieve fine'. Its underlying form has two identical fricatives, one before and one after the word boundary, i.e. /ze:f#fɛin/. In the surface representation, however, only one boundary consonant remains, i.e. [ze:fɛin].

Due to degemination, phonetic ambiguities may occur, which means that a speech fragment may correspond to more than one lexical entry, as is the case in our example. The utterance [ze:fɛin] does not only correspond to the two-word phrase *zeef fijn*, but may also be interpreted as a two-word phrase with a single underlying boundary consonant, namely as *zee fijn* 'sea fine' (/ze:#fɛin/). Thus, the first word of this two-word phrase may correspond to the word *zeef* as well as to the word *zee*.

Since listeners in an everyday listening situation are usually unaware of such phonetic ambiguities, the question arises whether degemination always occurs and, if so, whether this process is complete. In order to answer this question, Martens and Quené [1] recorded minimal pairs of two-word phrases of the type *zee fijn* versus *zeef fijn*, read at different speaking rates, that is: slow, normal and fast. Duration measurements showed longer boundary consonants for two-word phrases with underlying double consonants. They obtained this result even in the fast speaking rate, although the difference was not significant in this condition. Thus, Martens and Quené concluded that degemination is a gradual process, and that complete deletion of one of the members of the two consonants does not occur.

It may be argued, however, that degemination is complete if listeners can no longer differentiate between an underlying double or single consonant, despite the durational differences in the acoustic domain. This is true if correctly perceived segmentation does not exceed chance level. It was shown by Menert [2] that degemination and assimilation within words were perceptually complete at a fast speaking rate. However, Quené [3], using other types of ambiguities occurring at word boundaries, found that listeners are quite capable of using durational information in the boundary consonant and its surrounding vowels to segment correctly the two-word phrases used. Thus, durational differences in word boundary ambiguities can be a strong cue for correct segmentation.

In the present experiment we will, therefore, raise the question of whether listeners are able to use the small durational differences between the boundary consonants in minimal pairs of two-word phrases of the type *zee fijn* versus *zeef fijn*, when pronounced at a fast speech rate. It has also been shown that acoustic cues in the surrounding vowels may contribute to correct segmentation ([3]), and it seems likely that boundary consonant duration can only be interpreted well in relation to the surrounding speech sounds. Therefore, the second question we will address is whether information from the surrounding vowels can also contribute to correct segmentation. We will try to answer these questions by using a pseudo-gating task, in which after each gate subjects have to indicate which member of the minimal pair has been uttered, the degeminated or the non-degeminated two-word phrase. Implications of the obtained results for word recognition will be discussed as well.

2. METHOD

2.1. Materials

The stimulus material consisted of ten minimal pairs of two-word phrases. The members of the minimal pairs differed in their underlying representation, having either a double or a single boundary consonant. An example is the two-word phrase *zee fijn* 'sea fine' versus *zeef fijn* 'sieve fine'. In the former case, the (single) boundary consonant always belongs to the second word only (/..V#CV../), whereas in the latter case this post-boundary consonant also occurs in pre-boundary position (/..VC_i#C_iV..). All phrases were noun-adjective combinations, consisting of two monosyllabic and monomorphemic words. The pivotal consonant was always a voiceless fricative.

The ten minimal pairs were embedded at the end of semantically neutral carrier sentences and read in a sound-treated booth by a male native speaker of Dutch. The speaker was asked to read the sentences at a relatively fast speech rate. Stimulus sentences were mixed with other types of sentences to lead the speaker away from the goal of the present experiment.

The stimuli were presented in three different gates. The first gate included the sentence and the first vowel of the stimulus phrases, the second gate consisted of the sentence up to and including the ambiguous boundary consonant, and the third gate contained the sentence and the whole two-word phrase. The place of the gates was auditorily and visually determined with the aid of the oscillogram and a spectrogram. Gates were cut off at zero crossings to avoid clicks.

2.2. Subjects and Procedure

Subjects were 40 native speakers of Dutch, who participated voluntarily. They did not have any self-reported hearing impairment and were not aware of the goals of the experiment. The subjects were divided into two groups, each group receiving just one member of each minimal pair. In this way, effects of priming were avoided.

Stimulus phrases were presented in a random order. All three versions of the same stimulus phrase were presented consecutively over closed headphones. Subjects were tested individually in a sound-treated booth. They were asked to listen carefully to each stimulus and to indicate in a binary forced choice task whether they had heard a two-word phrase with a single or a double consonant, by circling that two-word phrase on answering sheets.

3. RESULTS

If subjects are able to use the durational cues present in the signal, we expect correctly perceived segmentation above chance level at the third gate, in which the whole two-word phrase was presented. If this is true, we can then deduce from responses at each gate whether durational information from the first vowel, the boundary consonant or the second vowel was the most important cue for their decisions.

At the first gate, the sentence including the vowel of the first word was presented. If subjects are able to segment the two-word phrases correctly as early as this first gate, then the onset and vowel preceding the boundary consonant must have provided sufficient cues. If correct segmentation is above chance level at the second, but not at the first gate, then the durational information of the boundary consonant is the most relevant cue. But if subjects can only segment the two-word phrases correctly after the third gate, they probably need the whole utterance as a framework in which to interpret the durational cues. Of course, if our subjects cannot arrive at a correct segmentation after the third gate, the two-word phrases must have been truly phonetically ambiguous.

Table 1: Number of correct segmentations broken down by the different two-word phrases and the three gates. N per two-word phrase per gate is 20. N over all two-word phrases with a single or a double consonant is 200.

two-word phrase	*gate 1*	*gate 2*	*gate 3*
zee fijn 'sea fine'	18	5	14
zeef fijn 'sieve fine'	1	15	10
ei slecht 'egg bad'	19	5	12
ijs slecht 'ice bad'	3	13	9
prei slecht 'leek bad'	15	3	10
prijs slecht 'price bad'	0	17	7
brie fijn 'brie fine'	17	1	10
brief fijn 'letter nice'	3	20	17
wei fraai 'meadow beautiful'	13	1	7
wijf fraai 'wife beautiful'	6	20	14
reu slim 'dog clever'	18	3	9
reus slim 'giant clever'	2	17	12
wee sterk 'woe bad'	20	9	15
wees sterk 'orphan bad'	0	10	5
moe slecht 'mom bad'	20	7	12
moes slecht 'pulp bad'	1	19	9
lei fraai 'slate beautiful'	15	3	8
lijf fraai 'body beautiful'	4	16	13
ree snel 'roe quick'	15	2	6
race snel 'race quick'	3	16	9
total single consonant	170	39	103
total double consonant	23	163	105

With the aid of the binomial distribution (p=0.5, N=20) it was determined that 15 or more correct responses per two-word phrase were required for correct segmentation to exceed chance level. Table 1 shows the number of correct segmentations per two-word phrase, and the total number of correct responses for all two-word phrases with single and double underlying boundary consonants. A Kolmogorov-Smirnov test was used to determine whether the overall percentage of correct segmentations was significantly above chance.

The results show no significant correct segmentation at the third gate. Our subjects segmented the two-word phrases correctly in only 52% of all cases. The Kolmogorov-Smirnov test showed no significant difference between the binomial and the observed distribution ($p = 0.38$), indicating that the observed distribution does not differ significantly from chance. The results for each two-word phrase showed that only two phrases were segmented correctly above chance. However, the partners of these two-word phrases were not segmented correctly. The partner of one of these two-word phrases was even segmented incorrectly above chance level.

When looking at gates 1 and 2, we can find more evidence that for correct segmentation our subjects were unable to use the durational cues in the signal. At gate 1 a strong preference was found for the two-word phrases with underlying single consonants. Almost all two-word phrases with underlying single consonants were segmented correctly above chance level, whereas their partners were segmented incorrectly. This probably is a direct result of the absence of a trace of the boundary consonant at that point in time. Thus, the vowel preceding the boundary consonant did not provide sufficient cues for correct segmentation.

At gate 2 this pattern was reversed. Here, our subjects indicated that they had heard a two-word phrase with a double boundary consonant in most cases, independently of which two-word phrase was actually read by the speaker. Thus, the presence of the boundary consonant did not provide sufficient cues to segment the two-word phrases correctly. Instead, subjects indicated that they had heard a first word consisting of a closed syllable.

Although the total number of correct segmentations was equal for two-word phrases with single or double consonants, results for each minimal pair seem to show a preference for one member at the third gate. This suggests that our subjects used other criteria when making a choice. Therefore, we determined whether there was a significant relationship between the percentage of recognised two-word phrases with a single boundary consonant and the relative word frequency of the first word (e.g. *zee* 'sea'). The relative frequency is defined as the absolute frequency of word A divided by the frequency of word A plus word B. Furthermore, the correlation between the percentage of recognised two-word phrases with a single consonant and the preferences of the subjects for one member of each minimal pair was determined. Preferences were obtained in a pen and paper test, in which subjects were asked to choose which one of the two members of each minimal pair was the best continuation of a semantically neutral sentence.

Results show no significant relation between preference and percentage of recognised phrases with a single consonant, although a trend in the expected direction was obtained ($r=0.327$, $p=0.35$). Thus, word preferences do not seem to have influenced the responses significantly at the third gate. The relationship between percentage of non-degeminated responses and frequency was significant ($r=0.70$, $p < 0.05$). Since effects of non-sensory information like word frequency are probably stronger when the acoustic boundary markers do not provide sufficient information, the significant effect of word frequency may indicate once again that the two-word phrases were perceptually ambiguous to our listeners.

4. DISCUSSION

We expected that if listeners were able to distinguish between read versions of a two-word phrase with an underlying single or double consonant, they would segment the two-word phrases correctly above chance. Otherwise, the two-word phrases would be phonetically ambiguous, indicating that degemination was perceptually complete. Furthermore, results from gate 1, in which the two-word phrase up to the first vowel was presented, and gate 2, in which the boundary consonant was added, would give more detailed information as to whether the durational and other acoustic cues from the first vowel and the consonant provide sufficient information for correct segmentation.

The results of the present experiment show that listeners cannot make a perceptual distinction between two-word phrases with underlying single or double boundary consonants. The mean percentage of correct segmentations does not exceed chance level, indicating that our subjects applied a guessing strategy. Hence, the degemination process produces a perceptual ambiguity, which means that degemination is complete at a high speech rate. Dupuis [4], in a similar experiment, obtained results showing a bias towards the two-word phrase with the closest correspondence to the surface representation, i.e. the two-word phrase with a single underlying consonant. A possible cause of this difference is that in his experiment subjects could freely choose a word, whereas in our experiment the two possibilities were already given.

The results at gates 1 and 2 again show that listeners are incapable of using the durational or other acoustic cues to segment the two-word phrases correctly. At the first gate, subjects indicated that they had heard the two-word phrase with the single boundary consonant, which contains a first word consisting of an open syllable. This is not surprising, since at that gate an open word, for instance *zee* 'sea', was presented to the subjects, followed by a silence. In terms of a theory of word recognition, for instance the cohort model ([5], [6], [7]) the open word *zee* 'sea' in the two-word phrase *zee(f) fijn* 'sea/sieve fine' is highly activated at gate 1. Therefore, subjects will tend to recognise the word *zee* and indicate that they have heard the two-word phrase with a single consonant in which *zee* is embedded.

At the second gate, however, subjects switch to a two-word phrase with an underlying double consonant. At that moment a closed word, for instance *zeef*, was presented completely, followed by a silence. At that moment the closed word (*zeef*) was activated highly, since more acoustic evidence was present in favour of the closed word *zeef* than in favour of the open word *zee*.

Thus, the listeners now prefer a two-word phrase with a double consonant containing the word *zeef.*

It could of course be argued that this result would have been obtained even if sufficient information had been present in the acoustic signal. Higher activation of a first word consisting of an open syllable at gate 1 and of a first word consisting of a closed syllable at gate 2 may have been sufficiently strong to overrule the detailed acoustic information. However, prior to the experiment the subjects were informed that incomplete parts of two-word phrases would be presented, and they were instructed to listen carefully. Furthermore, the results of the third gate clearly show a guessing strategy in our subjects, indicating that the small durational cues were of no help to our subjects. After the third gate, both *zee* and *zeef* probably have an equal activation, making a decision more difficult. And since our subjects were forced to choose at some point, they had to take a guess.

When the acoustic-phonetic information is not sufficient for correct segmentation, we would expect other types of, non-sensory, information to influence responses. Since no semantically biasing sentence context was present, the only types of non-sensory information could be the relative frequency of the first word in the two-word phrases and any preference for one member of each minimal pair. Preferences did not seem to affect the responses significantly. Word frequency, however, did influence the responses at the third gate.

It has been shown that non-sensory information, like for instance word frequency, becomes more important when acoustic information is less clear ([8]). Therefore, the observed effect of word frequency may indicate again that the durational and other acoustic cues in the signal were insufficiently clear, so that our subjects used other kinds of information in order to make a choice.

The finding that degemination is a perceptually complete process, resulting in phonetic ambiguities, has implications for word recognition in everyday communication situations. The observation that we are mostly unaware of the occurrence of such phonetic ambiguities is probably a result of there being sufficient semantic and pragmatic information to solve the ambiguity even before we are aware of it. If no disambiguating information is present, and if the speaker is aware of a possible ambiguity, he may solve it beforehand by inserting a pause between the two words.

However, if all disambiguating information is absent when a phonetically ambiguous two-word phrase is uttered, then recognition of either *zee* or *zeef* must be delayed. In that case, there is no left to right recognition of the words in the sentence and a choice may only be made on the basis of subsequent context ([9], [10]).

5. CONCLUSION

Our results show that the phonological process of degemination is in fact perceptually complete at word boundaries and at a high speach rate. This result is in accordance with Menert's ([2]) findings for assimilation and degemination within words. As a result of degemination, phonetic ambiguities actually do arise, which may cause problems for word recognition.

6. REFERENCES

[1] L. Martens and H. Quené, "Degemination of Dutch fricatives", In: R. Bok-Bennema and C. Cremers (eds.) *Linguistics in the Netherlands 1994*, pp 119-126, 1994

[2] L. Menert, "*Experiments on voice assimilation in Dutch: Prosodic structures and tempo*", Doctoral Dissertation, University of Utrecht, 1994.

[3] H. Quené, "Durational cues for word segmentation in Dutch", *Journal of Phonetics*, 20, pp 331-350, 1993.

[4] M.Ch. Dupuis, *"Perceptual effects of phonetic and phonological accomodation"*, Doctoral Dissertation, Univerity of Leyden, chapter 6, pp 83-103, 1988.

[5] U.H. Frauenfelder and L.K. Tyler, "The process of spoken word recognition", *Cognition*, 25, pp 1-20, 1987.

[6] W.D. Marslen-Wilson, "Parallel processing in spoken word recognition", *Cognition*, 25, pp 71-102, 1987.

[7] W.D. Marslen-Wilson and A. Welsh, "Processing interactions and lexical access during word recognition in continuous speech", *Cognitive Psychology*, 10, pp 29-63, 1978.

[8] H.B. Savin, "Word-frequency effect and errors in the perception of speech" *Journal of the Acoustical Society of America*, 35, pp 200-206, 1963.

[9] E.G. Bard, R.C. Shillcock and G.T.M. Altmann, "The recognition of word after their acoustic offsets in spontaneous speech: Effects of subsequent context", *Perception and Psychophysics*, 44, pp 395-408, 1988.

[10] C.M. Connine, D.G. Blasko and M. Hall, "Effects of subsequent sentence context in auditory word recognition: Temporal and linguistic constraints", *Journal of Memory and Language*, 30, pp234-250, 1992.