ISCA Archive
http://www.isca-speech.org/archive

ISCA ITRW on
Spoken Word Access Processes (SWAP)
Nijmegen, The Netherlands
May 29-31, 2000

DOES MORPHOLOGICAL INFORMATION INFLUENCE PHONETIC CATEGORIZATION?

K. Mauth
Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands

ABSTRACT

Two phonetic categorization experiments investigated the influence of sentence context on the perception of the inflectional morpheme *–t* (3rd person singular) in Dutch. The first experiment showed that people tend to give reliably more [t]-responses to ambiguous sounds when this phoneme is a morphologically meaningful unit that fits the context. A second experiment showed that this advantage is dependent on successful decomposition of the [t]-final string into a meaningful stem plus affix combination.

1. INTRODUCTION

As shown in the literature, phonetic categorization can be influenced by various sorts of information. *Lexicality* is well known to play an important role in subjects' judgments about ambiguous sounds at the beginnings (Ganong, [2]) or ends of words (McQueen, [4]). That is, on a word-nonword continuum people tend to label ambiguous sounds so as to form words.

Sentential *semantics* are also reported to modulate listeners' perceptions of ambiguous sounds. Connine [1] found that subjects were influenced by the semantics of the preceding context in the way they chose to label ambiguous sounds at the beginnings of words that were embedded in those sentences. On word-word continua listeners tended to label the ambiguous sounds such that the resulting word fitted the context semantically.

A very similar effect was observed by Isenberg, Walker and Ryder [3] who investigated the influence of preceding *syntactic* information on phonetic categorization. Listeners had to identify the function words *to* or *the* that preceded either the noun *gold* or the verb *go.* Depending on the following word people tended to give more syntactically congruent responses in the ambiguous region of the *to-the* continuum.

The present study sought to investigate how far preceding sentential context can influence the perception of inflectional morphemes. The relevant morpheme was the verbal 3rd person singular marker *–t* in Dutch. Listeners were presented with two different types of sentence-final words: inflected verbs (e.g., gaat, *leaves*) and nouns (e.g., plaat, *record*). In all conditions the final consonant was a stop plosive that varied in place of articulation along a continuum from [t] to [k] in which the [k] endpoints always formed nonwords. The main question was whether categorization of the final [t]-[k] continuum would be any different for verbs than for nouns. Is peoples' perception influenced by the fact that in the verb the very last consonant [t] constitutes an inflectional morpheme, which is predictable from the context? If so, this might be an indication of a decompositional process being at work.

Because the lexical shift that was expected for both sentence constructions where the [t] endpoints were real words (gaat vs. plaat) might mask a morphological effect, an additional nonword condition was included in the first experiment. In this condition both endpoints formed nonwords so that no bias towards one or the other endpoint would be expected.

2. EXPERIMENT 1

2.1. Methods

2.1.1. Materials

Two different types of preceding sentential context were constructed. One type ended with a verb (A: VP) while the other ended with a noun (B: NP). Because Dutch is a "verb second" language, questions had to be constructed that allowed for verb-final presentation.

(A: VP) Vraag jij of Jan morgen gaat?
Are you asking whether Jan leaves tomorrow?

(B: NP) Zie jij nog wel eens een plaat?
Do you see a record now and then?

Two sets of final items were chosen: one consisted of monosyllabic and the other of bisyllabic items. Within each set the three different types of words (verb, noun and nonword) were matched phonologically. Each nonword was presented in both the verbal and the nominal contexts.

To increase the variation of items, and by that increase the subjects' attention to the sentence contexts, 8 filler sentences were created. Half of these ended with a verb and half ended with a noun. Each sentence was also presented with a phonologically matched sentence-final nonword.

2.1.2. Stimulus construction

All sentences were recorded by a female native speaker of Dutch. The stimuli were recorded onto DAT tape and digitized afterwards onto a computer. These natural tokens were edited with the Xwaves waveform editor. An eight-step place of articulation continuum was constructed. A [t] and a [k] were spliced out of a natural waveform from one of the experimental sentences, with the cut being made at a zero crossing. Both stimuli were spliced such that they were 212 ms long. The six intermediate stimuli were constructed using a procedure developed by Stevenson [8] and Repp [7]. The mplitudes of the two waveforms were added sample by sample in different proportions. The vowel preceding the last sound was kept constant within each item set. In order to keep the transitions in the preceding vowel constant so that any bias towards [t] or [k] would be the same for each word or nonword the same vowels were used in all contexts. The vowels were taken from velar contexts (e.g., gaak), so had transitions towards a [k]. Any bias based on these transitions would therefore work against the predictions. The eight tokens from the continuum were then spliced onto all contexts. Each subject heard each step of each sentence 12 times, which resulted in a total number of 768 experimental sentences. Each step of the continuum for each of the filler sentences was only presented once which summed up to a total number of 128 filler presentations. The interval between items was 2500 ms.

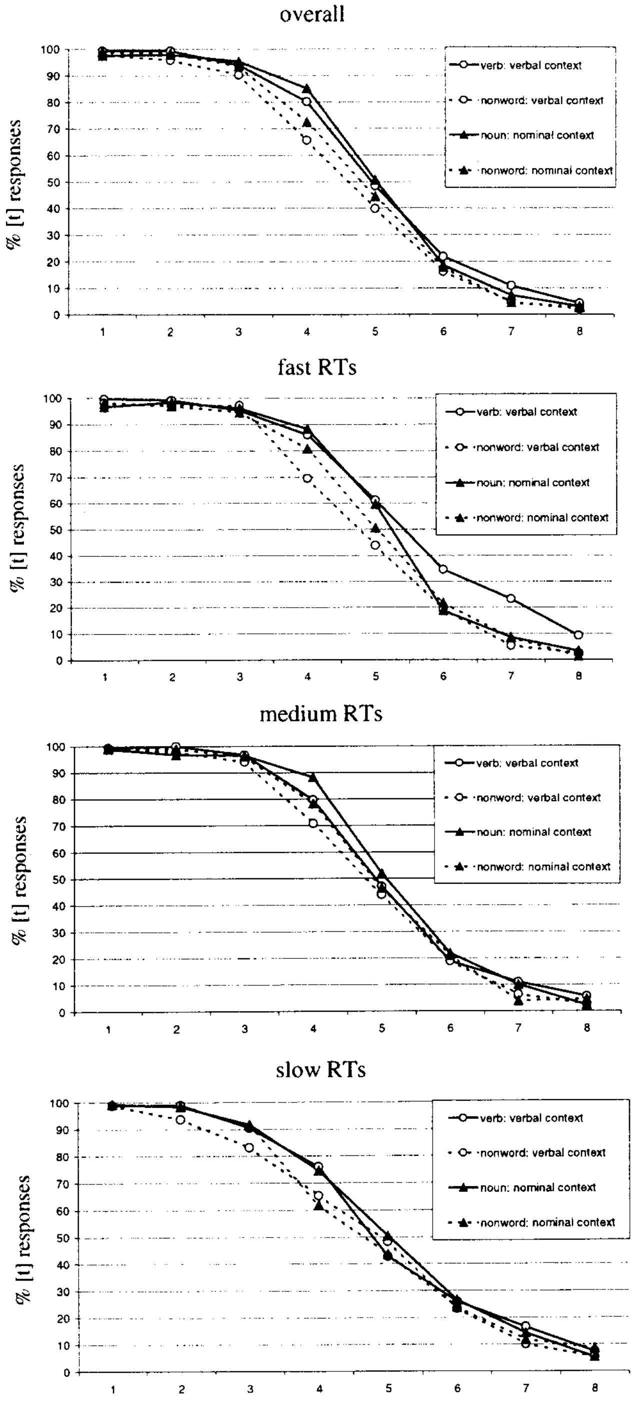

Figure 1. Results from Experiment 1: Percentage [t]-responses for each continuum step as a function of sentence type, overall, and in each RT range.

2.1.3. Subjects

Twenty student volunteers (15 female and 5 male) from the Katholieke Universiteit Nijmegen were paid for their participation. No subject reported any hearing impairment.

2.1.4. Procedure

Subjects were tested in groups of one to three in sound-proofed booths. The stimuli were presented via headphones at a comfortable listening level. Because of the large number of stimuli, subjects were tested twice, receiving 448 trials in each session. In the second session, reaction boxes were turned around so that reactions with each listener's preferred hand were evenly distributed over both [t]- and [k]-reactions. Subjects were instructed to give their responses as fast and as accurately as possible even if they were uncertain about the sound. In each part of the experiment there was a practice session of 24 sentences before the real experiment started. Items were presented in three blocks within each part so that subjects were allowed two breaks. Each part of the experiment lasted 50 minutes including breaks.

2.2. Results and Discussion

The percentage of [t]-responses was computed for each subject as a function of sentence type and stimulus continuum. In Figure 1 the categorization functions of the four different sentence types across item sets are plotted. Because a lexical shift typically shows up over a range of steps rather than at a single cross-over point (see also Pitt and Samuel, [6]), a boundary region was chosen. The boundary region was defined as extending from step 4 to step 7. Analyses were carried out on the proportion of [t] responses in this boundary region. A one-way repeated measures ANOVA showed an effect of lexicality (whether the last item was a real word or a nonword), $F(1,19)=35$, $p < .001$. The factor context type (whether the context predicted a noun or a verb) was not significant and there was no interaction between the two factors. Post-hoc t-tests showed that both the real verbs and the real nouns received significantly more [t]-responses than their nonword counterparts, $t(19)=4.59$, $p < .001$ (verb vs. nonword in verbal context) and $t(19)=4.43$, $p < .001$ (noun vs. nonword in nominal context).

Following Miller and Dexter [5], each listener's responses to each step in each sentence along the stimulus continuum were ranked and divided into three different reaction time groups: fast, medium and slow

RTs. Mean RTs for these groups were respectively 417 ms (SD = 126 ms), 545 ms (SD = 156 ms) and 779 ms (SD = 299). The percentage of [t]-responses to each stimulus in each RT range was then calculated for each subject as a function of context bias. For each RT range, one-way repeated measures ANOVAs were performed seperately on the proportion of [t] responses in the boundary region. In the fast RT range there was a significant effect of lexicality, $F(1,19)=27.82$, $p < .001$, and a significant interaction of lexicality with the factor context type, $F(1,19)=8.15$, $p < .01$. A post-hoc t-test showed a significant difference between real verbs and their nonword counterparts, $t(19)=5.98$, $p < .001$, while there was no such effect for the nominal item set.

The lexicality effect was still significant in the medium RT range, $F(1,19)=11.27$, $p < .01$, but the interaction between lexicality and context type vanished. A post-hoc t-test showed that the overall lexicality effect was due to a significant difference between real nouns and their nonword counterparts, $t(19)=2.37$, $p < .05$. There was no significant difference between real verbs and the relevant nonword condition.

In the slow RT range the lexicality effect disappeared and also pairwise post-hoc comparisons gave no significant effects.

In sum, there are clear indications that the functions for real verbs and nouns as compared to the relevant nonword conditions show different patterns over time. The lexicality effect for verbs is strongest in the fast RT range and gets weaker over time while the lexicality effect for nouns is only significant in the medium RT range.

3. EXPERIMENT 2

In Experiment 2, the *word* conditions were substituted by *possible word* conditions in order to look at the decompositional issue in more detail. These nonwords had a real noun embedded within them (vla *custard* + t = vlaat). Would listeners be more willing to label an ambiguous sound as [t] when, on the basis of the context (Vraag jij of Jan morgen vlaat?), they can decompose the last nonword into two meaningful units than when this decomposition would not make any sense (like in: Zie jij nog wel eens een vlaat?)? As in the first experiment those possible words were compared with nonword conditions.

3.1 Methods

3.1.1. Materials
The same preceding contexts were used as in Experiment 1. This time the *real word* conditions were replaced by the *possible word* conditions (C) and (D).

(C) Vraag jij of Jan morgen vlaat?
Are you asking whether Jan custards tomorrow?

(D) Zie jij nog wel eens een vlaat?
Do you now and then see a custards?

3.1.2. Subjects
Twenty student volunteers (15 female and 5 male) from the Katholieke Universiteit Nijmegen were paid for their participation. No subject reported any hearing impairment.

3.1.3. Procedure
Procedures were the same as in Experiment 1.

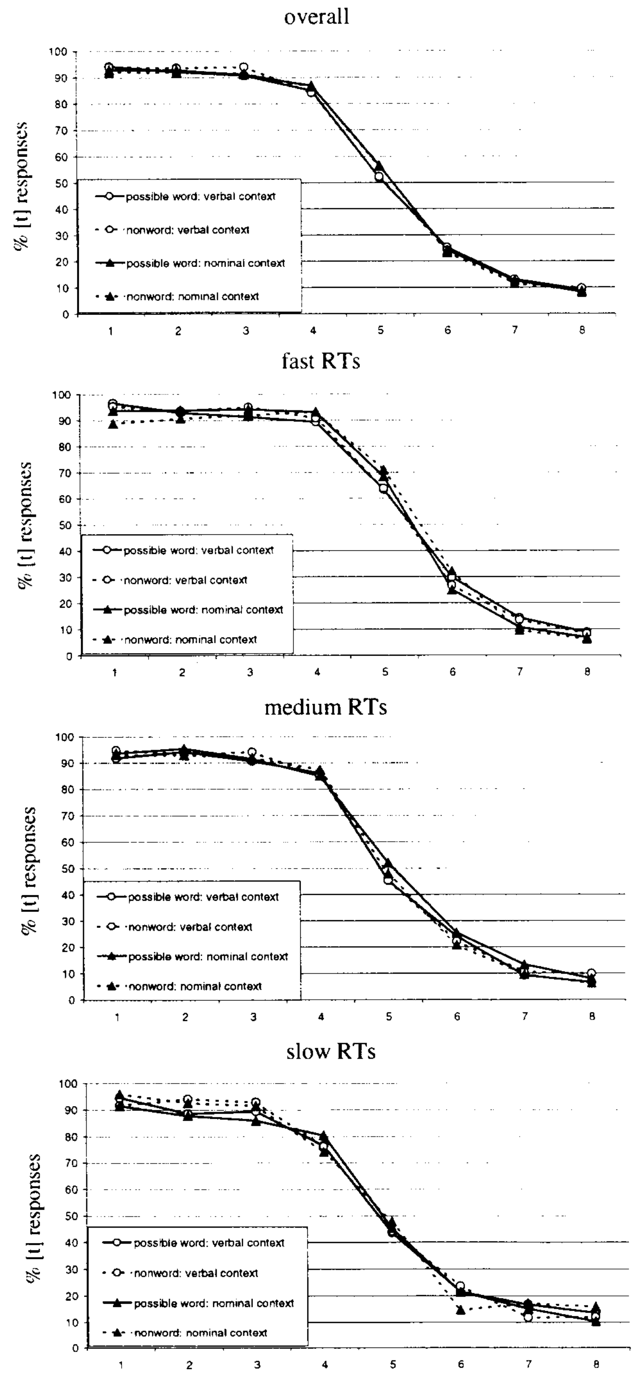

Figure 2. Results from Experiment 2: Percentage [t]-responses for each continuum step as a function of sentence type, overall, and in each RT range.

3.2. Results and Discussion

Again the percentage of [t]-responses was computed for each subject as a function of sentence type and stimulus continuum. The boundary region was defined as in the previous experiment from step 4 to step 7. A one-way repeated measures ANOVA on the proportion of [t] responses in the boundary region showed no lexicality effect, no context type effect and no interaction of the two factors. Again each listener's responses to each step in each sentence along the continuum were ranked and divided into three different RT groups. Mean RTs for these groups were respectively 427 ms (SD = 102 ms), 574 ms (SD = 132 ms) and 840 ms (SD = 315 ms). For each RT range one-way repeated measures ANOVAs were calculated. In the fast RT range there was a significant interaction of lexicality with context type, $F(1,19)=5.04$, $p < .05$. Although post-hoc pairwise comparisons revealed no significant effects, some of the differences were near significance: this might explain the overall interaction of lexicality with context type. Neither in the medium nor in the slow RT range were any of the main effects significant.

In sum, no clear difference could be observed between the possible words and the nonwords. If people did decompose the string into the noun and the verbal inflectional marker *–t* they did not use this decomposition to influence their categorization decisions. It would appear that for there to be a bias towards [t] responses, there needs to be a meaningful grammatical relationship between the two decomposed parts.

4. GENERAL DISCUSSION

The different patterns of the categorization functions for the verbal and nominal conditions in the first experiment suggest that listeners are sensitive to morphological information during phonological decision making. Especially when responding fast they seem to profit from the fact that the phoneme they have to judge has a morphological status. The second experiment, though, shows that the morphological status of the last phoneme on its own can not account for the effect found in the first experiment. Decomposition into two meaningful units that otherwise do not have a grammatical relationship does not seem to give people an advantage at that stage of processing.

Furthermore, the time course of the categorization bias also implies that the influence of the morphological information on phonetic categorization is temporally restricted. Only when responding fast did people benefit from this source of information. Apparently the decompositional process is a very rapid one that does not leave the morphological information available for perceptual decisions once it is integrated into the sentence context.

But that leaves open the question how far the obtained effect of the first experiment is actually dependent on the preceding sentential context. Would the same effect show up if listeners were not provided with any predictive information? This question has to be investigated in an additional experiment where the sentence-final words would be presented in isolation.

Although Experiment 2 suggests that possible words which have real words embedded within them are not treated differently from nonwords in terms of decomposition, it is evident from advertisements, for example, that people are very well able to decompose nonsense strings into parts that normally do not share a grammatical relationship. But because the newly formed words in this experiment do not have an "established" status and because the morphological influence only shows up at an early period of phonetic decision making, this ability to decompose nonsense strings might not have been of any advantage in the second experiment. Probably the current pattern would change if listeners were provided with a longer context in which the nonsense word could be established as a novel verb before an experiment.

5. ACKNOWLEDGEMENTS

Many thanks to James McQueen for his supervision and helpful discussions. Address for correspondence: Kerstin Mauth, Max-Planck-Institute for Psycholinguistics, PO Box 310, 6500 AH Nijmegen, The Netherlands. Email: kerstin.mauth@mpi.nl

6. REFERENCES

[1] Connine, C.M. (1987). Constraints on interactive processes in auditory word recognition: The role of sentence context. *Journal of Memory and Language*, **26**, 527-538.

[2] Ganong, W.F. (1980). Phonetic categorization in auditory word perception. *Journal of Experimental Psychology: Human Perception and Performance*, **6**, 110-125.

[3] Isenberg, D., Walker, E. & Ryder, J. (1980). A top-down effect on the identification of function words. Paper presented at the meeting of Acoustical Society of America, Los Angeles, California.

[4] McQueen, J.M. (1991). The influence of the lexicon on phonetic categorization: Stimulus quality in word-final ambiguity. *Journal of Experimental Psychology: Human Perception and Performance*, **17**, 433-443.

[5] Miller, J.L. & Dexter, E.R. (1988). Effects of speaking rate and lexical status on phonetic perception. *Journal of Experimental Psychology: Human Perception and Performance*, **14**, 369-378.

[6] 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.

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

[8] Stevenson, D.C. (1979). *Categorical perception and selective adaptation phenomena in speech*. Unpublished doctoral dissertation, University of Alberta, Edmonton, Alberta, Canada.