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

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

THE TIME-COURSE OF COMPETITION

A. H. van der Lugt
Exeter University, Exeter, United Kingdom

ABSTRACT

McQueen, Norris and Cutler ([1]) have shown that it is harder recognise a word when it is embedded at the beginning of a longer word because of competition between the embedded word and the longer word. In this study two identity priming experiments were aimed at replicating this competition effect and utilising it to investigate effects of mismatching information on word recognition. No evidence was found for competition between overlapping lexical candidates. The results are interpreted in the context of the SHORTLIST model ([2],[3]). Specifically, the absence of lexical inhibition is explained in terms of the time-course of the competition process.

1. INTRODUCTION

Like speech, spoken word recognition unfolds in time. The temporal nature of the recognition process poses an additional problem to the psycholinguist: when does an effect occur? In cross-modal form priming experiments, participants are presented with a visual target, to which a response (for example, lexical decision) is required, preceded by a spoken utterance that is structurally related to the visual target in the critical conditions. When visual target and auditory stimulus are related, the auditory stimulus is intended to facilitate or inhibit the processing of the visual target. This effect is measured with respect to a control condition in which the auditory and the visual stimulus are not related. A major advantage of this paradigm is the possibility to present the visual target at various times relative to the presentation of the auditory target. This allows one to actually trace the lexical activation of words in time.

Previous studies have shown that competition is a central mechanism of spoken word recognition. Using identity priming, Vroomen and De Gelder ([4]) showed that visual lexical decision to the Dutch word MELK (milk) was slower after hearing *melkaam* than after hearing *melkeum*. This difference was interpreted as evidence for competition. In Dutch, there are many words that start with *kaa(m)* but only few words that start with *keu(m)*. If the end of the speech input is consistent with many words, lexical decision to an overlapping competitor (e.g. melk) is facilitated to a smaller degree than when the final part is consistent with only a few words. Other evidence for lexical competition was obtained by McQueen, Norris and Cutler ([1]) for English using word spotting. They demonstrated that it is harder to spot a word like mess in *demess* which is the beginning of *domestic*, than in *nemess*, which is not the beginning of an English word. Similar materials were here evaluated in two identity priming experiments. The original aim of these experiments was to utilise the competition effect that was found by McQueen et al. to investigate effects of mismatching information on lexical access and competition.

2. EXPERIMENT 1

In Experiment 1 the participants listened to a list of bisyllabic nonsense items. Immediately after hearing the auditory stimulus two things could happen. In the first case a letter string, for example LES, was presented on a computer screen immediately after *choles* had been presented. In this case the participants were asked to decide whether the letter string was an existing Dutch word. In the second case, a warning signal was presented on the screen followed by a letter string, for example ZUIF after *zuifkaaf.* In this case the participants were asked to decide whether the string formed an adequate description of the beginning of the nonsense word that they had just heard. This 'catch' task was included to ensure that the participants were also attending to the beginnings of the auditory stimuli.

One would predict, in line with Vroomen en de Gelder's results, that visual lexical decisions on target words that are preceded by auditory stimuli which have those words embedded in them will be less facilitated when the auditory stimuli also form the onsets of longer words in Dutch. Competition with the longer candidate should lower the activation of the embedded word that is caused by the auditory stimulus, and thereby reduce the priming effect on the visual lexical decision on the embedded word. For instance, it should be harder to make a lexical decision to LES after hearing *choles* which is the beginning of *cholesterol*, than after hearing *boles*. The results from previous studies suggest that *boles* will mismatch with *cholesterol* to block activation of that word, so there will be no competition, and lexical decision to LES should be relatively easy.

2.1. Method

2.1.1. Materials

Thirty-two polysyllabic words with three or more syllables were selected from the Dutch CELEX lexicon. These words all satisfied the following constraints: neither the first syllable nor the first two syllables together formed an existing Dutch word, but the second syllable was a real word in Dutch. For instance, the first and the first two syllables of the Dutch word *cholesterol* (cholesterol), *cho* and *choles*, are not existing Dutch words, but the second syllable forms the Dutch word *les* (lesson). For every bisyllabic Word Onset (e.g. *choles*) two mismatching versions were created: a Minimal Mismatch condition and a Maximal Mismatch condition.

In the minimal change condition the initial consonant of the word onset was changed into a consonant that differed only in place of articulation (e.g. *foles*) in the maximal change condition the initial consonant was substituted by a consonant that differed in place, voicing and manner of articulation (e.g. *boles*). A set of 64 appropiate fillers was also made to avoid participants responding strategically to word onsets. These materials were used as the auditory primes in the present experiment.

The 32 experimental triplets were used to construct quadruplets of prime-target pairs. The embedded words of the triplets, for example *les* from *choles-foles-boles*, were used as visual targets in the three experimental conditions, Word Onset, Minimal Mismatch and Maximal Mismatch and in a control condition. In the control condition the response to the same visual target was measured when it was preceded by a different, unrelated, item from the stimulus set. Filler prime-target pairs were constructed in a similar fashion. The quadruplets were counterbalanced across four different versions of the experiment, so that each visual target appeared only once. Sixteen catch trials were equally spread across the list.

2.1.1. Subjects and procedure

Forty-eight subjects took part in Experiment 1 in four groups of 12 subjects for each version.

The stimulus presentation, the response timing and the data collection was performed by the NESU experimental system. Participants were seated in front of a computer screen. The auditory stimuli were presented binaurally via Sennheiser headphones. The visual stimuli were presented centered on the computer screen in white against a black background in 36 point Arial font for 1000 ms. The participants were instructed to listen to a list of nonsense bisyllables. At the offset of each auditory stimulus two things could happen on the computer screen. If a letter string was presented, they were asked to perform a lexical decision task on the visual stimuli. They were asked to press a green response button marked "JA" (yes) with their preferred hand as quickly as possible, whenever they saw an existing Dutch word on the screen. If the letter string that was presented was not a real word, they were asked to press a red button marked "NEE" (no) with their other hand. On the catch-trials a warning signal "***" was presented visually for one second starting at the offset of the prime. This warning signal was immediately followed by the visual presentation for one second of a letter string. In this case, the participants were asked to indicate by a button press whether the letter string formed an adequate description of the beginning of the auditory prime or not. The main experiment was preceded by a block of 9 practice trials.

2.2. Results

First, the performance on the catch trials was analysed separately to investigate if the participants were listening to the beginning of the auditory primes. All of the participants responded correctly to more than 80% of the catch trials, suggesting that they were not ignoring the first part of the primes in this experiment.

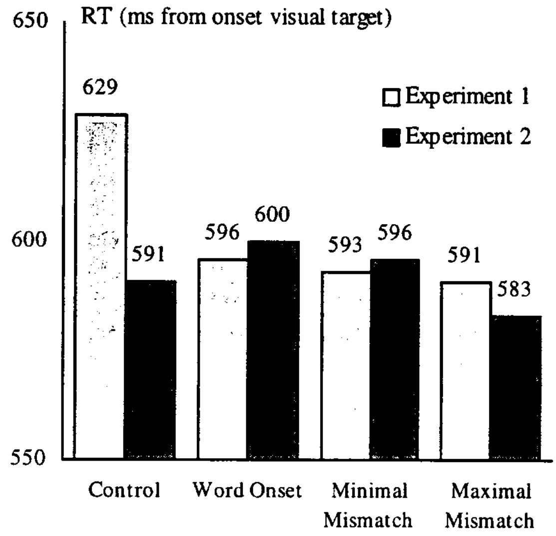

Figure 1 *Average lexical decision latencies for Experiment 1 and 2. Visual targets were presented at the offset of the embedded word in the auditory prime for Experiment 1 and at the onset of the embedded word for Experiment 2.*

Lexical decision latencies were measured from the onset of the visual targets on the screen. Analyses of variance (ANOVAs) were performed on these decision latencies and on the proportions of errors with both subjects ($F1$) and items ($F2$) as the repeated measure (see Figure 1 for the mean responses from the subjects analysis). Non-responses and responses longer than 1200 ms were treated as missing data. The missing latencies were substituted by the mean reaction time per subject per condition for the subjects analysis and by the mean reaction time per item per condition for the items analysis. Listeners were on average 35 ms faster in making lexical decisions to strings that were preceded by a related auditory prime, in comparison to an unrelated control. The differences between the three related conditions, Word Onset, Minimal Mismatch and Maximal Mismatch, were very small. The ANOVAs of the reaction times suggested that there was an overall effect of priming in the related conditions, $F1(3,132) = 6.28$, $p < .01$, $MSE = 2{,}457$; $F2(3,93) = 3.77$, $p < .05$, $MSE = 3{,}403$. Post-hoc Tukey HSD tests ($p < .05$) and pairwise comparisons confirmed that the Word Onset condition, the Minimal Mismatch and the Maximal Mismatch were not significantly different from each other ($p > .1$), but that all three related conditions were significantly different from the control condition by subjects and by items (Word Onset: $t1(47) = -2.48$, $p < .05$; $t2\ (31) = -2.25$, $p < .05$, Minimal Mismatch: $t1(47) = -3.29$, $p < .01$; $t2(31) = -2.77$, $p < .01$ and Maximal Mismatch: $t1(47) = -3.37$, $p < .01$; $t2(31) = -3.58$, $p < .01$). No reliable differences were found in the error data ($F1$ and $F2$, $p > .1$).

In summary, identical priming effects were found for the Word Onset, the Minimal Mismatch and the Maximal Mismatch condition when the visual target was presented at the offset of the embedded word. No differences were found between the three priming conditions. Lexical decision is equally facilitated by related auditory primes that form the beginning of a longer word as by primes that don't. The participants' performance on the secondary catch task suggests that this absence of a differential effect of competition can not be attributed to some kind of attentional strategy, where subjects are ignoring the first part of the auditory stimuli. To investigate whether differential competition effects arise at an earlier stage of processing, a second priming experiment was conducted.

3. EXPERIMENT 2

In Experiment 2 the visual stimuli, both for the experimental trials and the 'catch' trials, were presented at an earlier point relative to the auditory prime: at the onset of the embedded word.

3.1 Method

3.1.1. Materials

The materials were the same that were used in Experiment 1. For all of the auditory stimuli the onset of the second syllable was determined by visual and auditory inspection of the waveform. For this purpose the Xwaves speech editor was used. The order of presentation was the same as in Experiment 1.

3.1.2. Subjects and procedure

Forty-eight subjects (equally divided between 4 groups) participated in Experiment 2. Subjects that already participated in the previous experiment were excluded.

The procedure was similar to the procedure used in Experiment 1. The only difference was that the visual stimuli, both the letter strings on the identity priming trials and the warning signal on the catch trials, were now presented at the onset of the second syllable.

3.2. Results

Analyses of the performance on the secondary task showed that all subjects correctly responded in more than 80% of the catch trials. This suggests that they were indeed attending to the auditory stimuli.

Lexical decision latencies were measured from the onset of the visual targets on the computer screen. Responses later than 1200 ms and cases where subjects did not respond at all were treated as missing data. Missing data points were again replaced by the relevant mean (per subject or per item) per condition.

The differences between all four conditions, were very small (see Figure 1 for the mean results from the subjects analysis). The ANOVAs of the RTs suggested that there is no facilitatory effect of priming ($F1$ and $F2$, $p > .1$). Post-hoc Tukey HSD tests and pairwise comparisons confirmed that the Word Onset condition, the Minimal Mismatch and the Maximal Mismatch were not significantly different from the Control baseline and each other ($p > .1$). Also in the error data no reliable differences were found ($F1$ and $F2$, $p > .1$).

To sum up, in Experiment 2 no priming effects were found when the visual target was presented at the onset of the embedded word. In Experiment 1 equivalent priming was found in all three conditions when the visual target was presented at the offset of the embedded word. Unlike the earlier studies by Vroomen and de Gelder and McQueen and collaborators, the present results demonstrate no effects of competition between overlapping lexical candidates.

4. SIMULATIONS

A possible explanation of the absence of lexical inhibition could be given in terms of the time-course of the competition process. If visual processing starts at the onset of the embedded word in the auditory stimulus (Experiment 2), the embedded word (e.g. *les*) might not be activated in any of the three critical conditions and, therefore might not prime visual lexical decision. If visual processing starts at the offset of the embedded word (Experiment 1), the embedded word should be activated and prime visual lexical decision in all three conditions, however, the embedding word (e.g. *cholesterol*) could be activated to a similar degree in all three conditions.

This possibility was investigated by running simulations with SHORTLIST ([2],[3]) operating on a 20,000-word lexicon of Dutch. For the simulations, all the experimental items were transcribed and used as input to the model. Mid-class phoneme transcriptions (see also [2]) were added to the Dutch phonemic inventory in SHORTLIST, in order to model the difference between minimal and maximal consonantal changes. The performance of the model on the three different conditions was compared in time slices, one for each phoneme in the input, followed by four time slices of silence. The mean activation functions for the target words and the longer embedding words are shown in Figure 2.

The time course of the predicted competition process can be investigated by looking at the activations of the target words at different time slices. Up to the penultimate phoneme of the experimental item there are no differences in level of activation between items from the three conditions. At the offset of the target word, targets from the Word Onset condition have a much lower activation than targets from the Maximal Mismatch condition, and targets from the Minimal Mismatch condition seem to be an intermediate case. At the first silence there is hardly any difference in activation left between the Minimal and the Maximal Mismatch condition, but both are still more highly activated than targets from the Word Onset items. At later slices the three functions converge completely.

Experiment 2 could be tapping into the competition process too early (visual processing starting at about "-2" on Figure 2) such that no priming is found because the embedded target is not yet activated. Experiment 1 could be tapping too late (visual processing starting at about "P" on Figure 2). At this point in time the target word is activated and primes lexical decision in all three conditions, but in this case competition could have acted to rule out the longer candidate in all three conditions before the lexical decision is made.

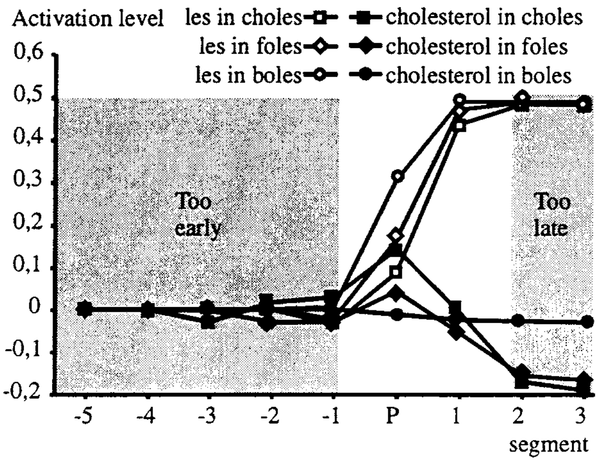

Figure 2 *Mean activation levels for the materials from Experiment 1 and 2 in SHORTLIST. Activation functions are shown for the embedded target words and for the competing longer candidates. The activation functions are aligned relative to the final phoneme of the embedded word ("P").*

5. DISCUSSION

In Experiment 1, with the visual target presented at the offset of the auditory stimulus, a facilitatory effect was observed for all the form related conditions. In Experiment 2, the visual target was presented at the beginning of the embedded word, and no priming was found for any of the conditions. In short, these experiments find no effect of competition with a longer overlapping candidate for recognition. This is somewhat unexpected given the results of McQueen et al., and those of Vroomen & de Gelder. The SHORTLIST simulations confirm the viability of an explanation of the results from the two form priming experiments in terms of the time course of competition. The first experiment probably taps in too late to observe competitioneffects and the second experiment taps in too early.

An obvious next experiment would be an identity priming experiment where the visual targets are presented about halfway through the embedded words in the auditory stimulus. However, there is an important problem with such an experiment: what is 'halfway'? The embedded target words were all monosyllabic, but they did not all have the same syllabic structure. For the targets (19) with a Consonant-Vowel-Consonant structure the midpoint could arguably be defined as the middle of the vocalic portion. The other 13 targets, however, had a Consonant-Vowel structure and it would be hard to establish where exactly 'halfway' is for these items. A different solution would be to just present the visual stimuli halfway in time. For Experiment 2 the durations of the embedded target words were measured in order to present the primes at the onset of the embedded word. These durations could be used to present the visual stimuli after half of the duration of the embedded word. However, this approach is even more problematic, since halfway in time is not necessarily halfway in information.

Although it is impossible to establish an unambiguous informational midpoint for the entire stimulus set in the present study, it would still be possible to hunt down the effect by presenting the visual stimuli at various points in time between the onset and the offset of the embedded word in the auditory stimulus. In the absence of any indication *when* the effect would be there, should there be any, this would be an almost impossible quest.

In the present study, and elsewhere, it is demonstrated that implemented models of spoken word recognition, like SHORTLIST, provide an invaluable tool to evaluate possible effects at different points in time. Such an evaluation allows the experimenter to interpret effects, and their absence, while considering the temporal dimension of spoken word recognition.

6. ACKNOWLEDGEMENTS

This research was carried out while the author was at the Max Planck Institute for Psycholinguistics in Nijmegen, The Netherlands, supported by a grant from the Max Planck Society in Munich, Germany. I am grateful to James McQueen and Anne Cutler for their excellent advice. I also thank Dennis Norris for making available the SHORTLIST model and Joost van de Weijer for his help in preparing the stimuli..

Correspondence should be addressed to A. H. van der Lugt, University of Exeter, Washington Singer Laboratories, Perry Road, Exeter, UK EX4 4QG. Email: A.H.Van-Der-Lugt@ex.ac.uk

7. REFERENCES

[1] McQueen, J. M., Norris, D. & Cutler, A. (1994). Competition in spoken word recognition: Spotting words in other words. *Journal of Experimental Psychology: Learning, Memory and Cognition*, **20**, 621-638.

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

[3] Norris, D., McQueen, J.M., Cutler, A. & Butterfield, S. (1997). The possible-word constraint in the segmentation of continuous speech. *Cognitive Psychology*, **34**, 191-243.

[4] 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.