Lexical effects in phoneme monitoring:
Facilitatory or inhibitory?

Uli H. Frauenfelder
Max-Planck-Institut für Psycholinguistik
Nijmegen, The Netherlands

Juan Segui
Laboratoire de Psychologie Experimentale
Université René Descartes and EPHE
Paris, France

Ton Dijkstra
Max-Planck-Institut für Psycholinguistik
Nijmegen, The Netherlands

Abstract

This paper addresses the questions of **how** and **when** lexical information influences phoneme detection in two phoneme monitoring experiments. In the first, the position of the stop consonant target (word initial, before uniqueness point, after uniqueness point, word final) and the lexical status of the target bearing item (word or nonword) were manipulated to pursue the temporal question. A contribution of the lexical level to phoneme detection (reflected by large RT differences between targets in words and nonwords derived from the words) was found only when the target came in the two positions after the uniqueness point. In a second experiment, the contribution of the lexicon was made incompatible with the bottom-up evidence for targets by placing them in words where they did not belong (p target substituted for l producing "stimupi"). No inhibitory effect of the lexical level was obtained even in cases where the target and substituted phonemes differed minimally. These results taken together indicate that the lexicon exerts its effect only **after** word recognition and as **positive** feedback suggesting strong limitations in the way in which lexical information can affect speech perception.

Introduction

No-one would dispute the claim that we recognize words on the basis of an analysis of the speech sounds of which they are composed. Controversial, at least in psycholinguistic circles, is the inverse claim that our perception of speech sounds depends upon the words they make up. In this paper, we will evaluate these claims about the relative importance of **bottom-up** and **top-down** processes mediating between the sublexical and lexical representations. To arrive at a proper description of the information flow between these two levels, we will address the questions of **how** and **when** the lexical and sublexical information sources are brought together.

To investigate these questions we used the phoneme monitoring task in which subjects are asked to detect as quickly as possible previously specified phoneme targets that appear in sentences or lists of words. Previous research [1,4,6,7] has shown that phoneme detection latencies are sensitive to lexical variables indicating an influence of the lexical level upon speech perception as reflected by the phoneme detection process. Our objective here is to examine empirically two opposed accounts of such lexical effects.

In **autonomous models** of language processing [3], it is assumed that bottom-up processes produce their output autonomously; top-down lexical information is not allowed to influence the bottom-up mechanisms responsible for phoneme perception. In order to account, nonetheless, for the presence of lexical effects in phoneme monitoring, "race models" [1,2] and "dual code models" [4] have been developed in which phoneme identification can be made on the basis of two different "outlets" or representations: a lexical and a sublexical level. In a race model account, there are two independent ways in which a phoneme target can be detected. The first target detection procedure depends upon the computation of a sublexical representation. In the second, target detection depends upon lexical access which makes available the phonological information associated with a particular accessed lexical entry. There is a race between these two processes with the one that reaches completion first providing the phoneme detection response. The presence or absence of lexical effects is explained in terms of the outcome of the race between these two independent and competing outlets.

Interactive activation models are designed to account for the integration of multiple sources of information or constraints in speech perception. The most explicit model constructed within this framework is TRACE [5]. In TRACE there are several levels of interconnected processing units corresponding to distinctive features, phonemes and words. The critical interactive aspect of this model is that word units can provide top-down feedback to phoneme units by increasing their level of activation. Hence, phoneme recognition (the moment a phoneme reaches a criterial level of activation with respect to the other phonemes) depends on **both** the amount of bottom-up activation from the distinctive feature level and the amount of top-down activation from the word level. Subjects responding in the phoneme monitoring task are assumed [8] to make direct and exclusive use of activated phoneme units. The presence or absence of lexical effects in phoneme monitoring is explained within the TRACE framework by varying lexical contributions to the phoneme's activation.

Although these two basic model types are radically different in nature, they make many of the same predictions and appear to be consistent with most of the data available in the phoneme monitoring literature. Given this state of affairs, it is critical to collect additional performance data that will allow us to further constrain these types of models. In particular, it is essential to determine how and when the lexical level contributes to the speech analysis as reflected by the phoneme monitoring task.

In order to trace the time-course of lexical effects, we selected targets in four different positions with respect to the uniqueness point (UP) of the word. The UP was defined as that point at which a word's initial part is shared by no other word listed in a phonetic dictionary. Nonwords were created from these target-bearing words by changing one or more phonemes, but keeping the target's local phonemic environment as constant as possible. The differences in detection times between phoneme targets in the same position in matched words and nonwords provided an approximate measure of the lexical contribution.

EXPERIMENT I

Subjects

Thirty-eight undergraduates at Nijmegen University, all native speakers of Dutch, were paid to participate in the experiment.

Materials and procedure

The test stimuli consisted of 120 words and 120 matched nonwords. The target phonemes (p , t , or k) occured in four different positions within target-bearing words (word onset, before the UP, after the UP and word offset) and nonwords (nonword onset, before the nonword point, after the nonword point and nonword offset). The nonword point (NWP) is that point at which the item becomes a nonword moving from its beginning towards the end.

Target-bearing items were embedded in counterbalanced lists made up of other words and nonwords not containing the target phoneme. Twelve such lists, each containing 60 items, were created and divided into two blocks for counterbalanced presentation to the subjects. For each list subjects were asked to detect as quickly as possible one of the three targets (specified by means of a visual display).

Results

Mean reaction times (measured from the burst of item-initial targets and from closure for the targets in the remaining three positions) were computed for each subject and each experimental item. All responses less than 100 ms. or greater than 1000 ms. were not included in the computation of the means. Three subjects with more than 15 % errors were also excluded from the analysis. Figure 1 shows the results for words and nonwords broken down according to target position.

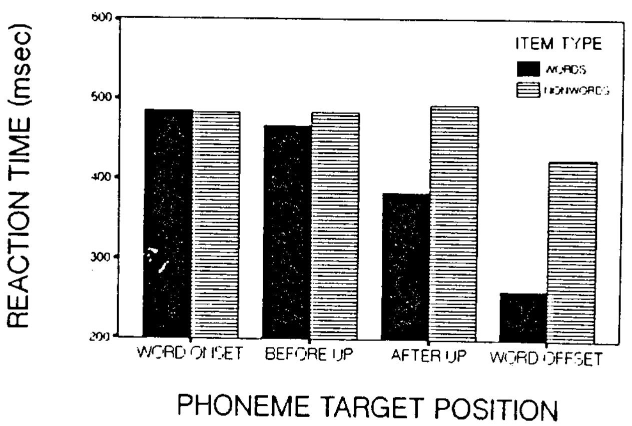

Fig. 1. Mean RTs for words and matched nonwords as a function of phoneme target position

An analysis of variance showed that the two main effects, lexical status and target position both were highly significant by subject ($F_1(1,34) = 371$; $p < .001$; $F_1(3,102) = 136$; $p < .001$) and by item ($F_2(1,9) = 48.2$, $p < .001$; $F_2(3,27)$, $p < .001$). The interaction between these two effects was also highly significant ($F(3,102) = 168$, $p < .001$). Post-hoc comparisons using the Scheffé test (S-method) revealed that the effect of lexical status was significant (at the .01 level) only for two target positions after UP.

Discussion

This experiment has shown a clear interaction between target position and lexical status of the target bearing item. The differences in detection latencies between targets in matched words and nonwords increased as the targets appeared later in the word. This increase is not linear; the differences for the two positions after the UP were significantly larger than those for the initial positions.

The global pattern of facilitation, in particular, the sharp increase in facilitation after the UP, is consistent with both the autonomous and interactive models. For the former, the lexical level can influence phoneme detection only after the phonological code associated with the target bearing word has been accessed. This code is normally assumed to become available at a discrete moment in time once the word has been recognized, that is, after the UP. In TRACE, the strength of lexical feed-back increases with the level of activation of the lexical units containing the target phoneme. Hence, the lexical level is involved from the beginning of the perceptual process, but its effects build up gradually and continuously as the lexical units themselves receive more activation. The activation of the target-bearing word should increase dramatically around the UP, and as a consequence so should the lexical influence on the phonemes.

In the following section we want to examine more closely **how** lexical information affects phoneme monitoring. In particular, we want to explore whether the lexical level can exert not only a **facilitatory** effect as in the first experiment but also an **inhibitory** effect upon the phoneme detection procedure. In order to test for this latter possibility, we created a situation in which the lexical information was made incompatible with the bottom-up evidence for the target. Subjects were asked to detect a target phoneme that appeared in the place of another phoneme situated after the UP (e.g. replacing the t in "simplicité" by the target phoneme d giving "simplicidé").

The two models described above appear to make different claims concerning the existence of inhibition effects. The interactive activation model predicts that a phoneme arriving after the UP (such as the t in simplicité) receives excitatory feed-back from the lexical level. This phoneme in turn inhibits the other phonemes (such as the substituted d) that occur at this point in the sequence. The decreased activation level of the target phonemes (d) should translate into slower detection times as compared with those to detect the same phoneme d in another nonword such as "fimplicidé" where the target phoneme should submit to neither excitatory nor inhibitory influences from the lexical or phoneme levels respectively. As a consequence, any difference between the detection times for the identical targets d in the same local phonemic environments " i **d** é " in the direction of slower detection times for the former type of nonword "simplicidé" would constitute evidence for mediated lexical inhibition effects.

The autonomous race model does not allow for inhibitory effects of this type. Since the two competing response outlets, lexical and pre-lexical, function independently, the lexical code cannot affect the elaboration of the pre-lexical code. Furthermore, the lexical code cannot contribute to the detection response for the target in "simplicidé" since this phoneme target is not contained in the lexical code for the word corresponding to the initial part of this nonword. As a consequence the target is always detected on the basis of the pre-lexical code for both nonwords leading the autonomous race model to predict no difference between the detection times in the two types of nonwords.

EXPERIMENT II

Subjects

Eighteen students of the University of Paris V, all native speakers of French, participated voluntarily in this experiment.

Materials and procedure

The test items consisted of 12 matched pairs of three- or four-syllable nonwords containing targets (4 pairs for each of the three targets: /d/, /t/, /k/). The inhibitory nonwords (INW) were constructed by replacing phonemes located after the UP in words by the target phoneme. Thus, for example, from the word "simplicité" whose UP lies well before the target, a INW item "simplicidé" was derived, in which the target-phoneme /d/ replaces the original phoneme /t/. Replaced and target phonemes differed only in the feature of voicing.

Matching neutral nonwords (NNW) were derived from each INW by replacing the initial phoneme of INW items by another phoneme of the same manner of articulation (e.g., from the INW "simplicidé" the NNW item "fimplicidé" was created). All nonword-points for NNW items were located before the end of the second syllable, well before the target phoneme.

Eighteen target-bearing words (six for each target type) were also included to confirm the existence of lexical facilitation effects; in half of these the target-phoneme was located before the UP (for example, "ouverture" ("opening") with the target-phoneme /t/), the other half with the target after the UP (for example, "profitable" with the target-phoneme /t/). These target-bearing words and nonwords were embedded in one of three experimental lists of 64 items each (32 words and 32 nonwords).

Results

Less than 5 % of the responses were eliminated for the computation of the means (latencies smaller than 100 ms. or longer than 1000 ms.). Figure 2 summarizes the means for both the words and the nonwords.

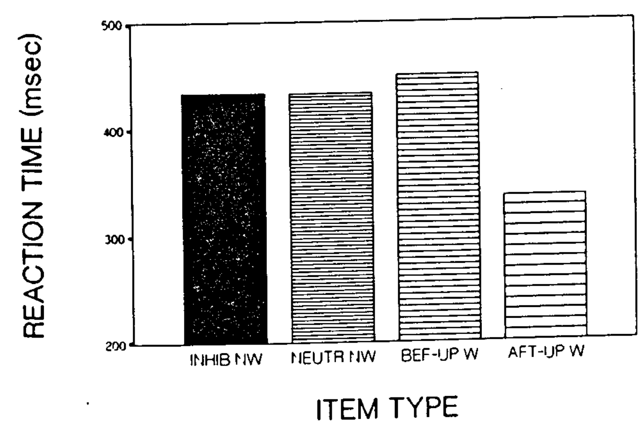

Fig. 2. Mean RTs for inhibitory and neutral nonwords and for words with target before and after UP

An analysis of variance performed on the nonword data indicated that neither the factor nonword type nor the type of target-phoneme (/d/, /t/, /k/) introduced significant effects ($F < 1$ in both cases). The interaction between the factors was also not significant ($F(2,34) = 2.16$, $p > .10$). Furthermore, there was no difference in the percentage of errors or omissions for the two nonword conditions. An analysis of variance for the words, however, showed that the difference between the two conditions (before and after UP) was highly significant ($F_1(1,17) = 55.40$, $p < .0005$).

Discussion

In this experiment, we have investigated whether the lexical influence, observed to be highly facilitatory after the UP in the first experiment, can also be inhibitory when the lexical information is incompatible with the target to be detected. The results provided no evidence for such inhibitory lexical effects, but did replicate facilitatory lexical effects. Predictions of inhibition are implicitly made in the interactive activation framework, but are excluded by the autonomous model. As a consequence these results appear to be more compatible with the autonomous model. The absence of inhibition, although problematic for interactive models can, nonetheless, be explained within this framework. According to such an account, the bottom-up activation of the target phoneme is so effective that the inhibitory influence of the replaced (appropriate) phoneme does not show up. The strength of the bottom-up activation dominates and hides the ephemeral inhibitory lexical effect. If this account is correct, then one might expect inhibitory effects to vary as a function of the strength of bottom-up activation. We are now in the process of exploring the possibility that inhibitory effects will emerge with acoustically less clear targets.

The two experiments taken together suggest an asymmetry in the way lexical information can contribute to the bottom-up analysis underlying phoneme detection. The results presented here indicate that the lexicon exerts a **facilitatory** but not an **inhibitory** influence upon bottom-up processing **after** word recognition. Thus, these results show strong limitations in the way in which lexical information can affect phoneme processing that must be taken into account by both interactive and autonomous models.

REFERENCES

1) Cutler, A., Mehler, J., Norris, D. & Segui, J. Phoneme identification and the lexicon. *Cognitive Psychology*, 1987, (in press)

2) Cutler, A. & Norris, D. Monitoring sentence comprehension. In W.E. Cooper & E.T.C. Walker (Eds.) *Sentence processing: Psycholinguistic studies presented to Merrill Garrett*, Hillsdale, N.J.:Lawrence Erlbaum Associates, 1979.

3) Forster, K.I. Levels of processing and the structure of the language processor. In W.E. Cooper & E.C.T. Walker (Eds.) *Sentence processing: Psycholinguistic studies presented to Merrill Garrett*, Hillsdale, N.J.: Lawrence Erlbaum Associates, 1979.

4) Foss, D.A. & Blank, M.A. Identifying the speech codes. *Cognitive Psychology*, 1980, 12, 1-31.

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

6) Marslen-Wilson, W.D. Function and process in spoken word recognition. In H. Bouma & D.G. Bouwhuis (Eds.) *Attention and Performance X: Control of Language Processes*. Hillsdale, N.J.: Lawrence Erlbaum Associates, 1984.

7) Segui, J., & Frauenfelder, U. The effect of lexical constraints upon speech perception. In F. Klix & H. Hagendorf (Eds.) *Human Memory and Cognitive Abilities: Symposium in Memoriam Hermann Ebbinghaus*. Amsterdam: North Holland, 1986.

8) Stemberger, J.P., Elman, J.L. & Haden, P. Interference between phonemes during phoneme monitoring: Evidence for an interactive activation model of speech perception. *Journal of Experimental Psychology: Human Perception and Performance*, 1985, 11(4), 475-489.