ASSIMILATION OF VOICE AND PERCEPTION OF VOICING: EFFECTS OF PHONETIC CONTEXT

I.H. Stins & R.J.K. van den Berg

Institute of Phonetics, University of Nijmegen
P.O. Box 9193, 6500 HD Nijmegen, The Netherlands

0. Abstract

The results of an earlier experiment contained indications that the perception of voicing in Dutch two-obstruent sequences was confounded by assimilation of voice. In this paper we will briefly discuss this study (section 2) and present the results of an experiment that was specifically designed to investigate such phonetic context effects on the perception of voicing in C1C2 sequences [5]. This experiment is presented in section 3 of this paper. In section 4, several hypotheses will be forwarded that may explain the results obtained. In order to be able to choose between these hypotheses a production experiment was run, which is discussed in section 4.

All experiments investigated homorganic two-obstruent sequences (C1C2) to avoid problems arising from the use of homorganic geminates. Because of restrictions inherent in Dutch [6] the sequences consisted of a phonologically voiceless-obstruent (C1) followed by a phonologically voiced one (C2).

2. Investigation of optimal stimulus form

In this first experiment [4] we investigated the perception of voicing in two-obstruent sequences that were part of two successive syllables (VC1V2V3C1C2). One of the aims of this study was to investigate whether the phonological status of the initial consonant (C1) of the first word pair in isolation and its effect on the perception of voicing in such sequences. To this purpose the C1C2 sequences were embedded in three types of linguistic context:

(a) a word pair that was part of a meaningful sentence

(b) the same word pair in isolation

(c) the same pairs as used in conditions (a) and (b) but in the initial position of a sentence

The consonantal sequence in which assimilation of voice was perceived as a sequence of two voiced (or two voiceless) obstruents was investigated by linguists who scored instances of assimilation by marking sentences in which the second obstruent had been assimilated to the first obstruent. However, one may perceive two consonants as having the same voicing status in spite of the fact that assimilation of voice did not take place [14]. In that case, it is obvious that voicing cues (i.e., acoustic cues to the voicing status of the consonants) other than the auditory result of the production or absence of vocal fold vibration will be used by the listener. At the Institute of Phonetics in Nijmegen (IPP) the effect of some of these cues on the perception of voicing in two-obstruent sequences was being investigated. In one of the experiments [13], a number of cues were found to affect the perception of voicing in such sequences [12, 13].

One of the effects that may affect the perception of voicing in C1C2 sequences appears to be the degree of similarity between the consonants in the context. Indications to that effect were found in the earlier study [4] set up to investigate the most suitable type of stimuli for a large series of experiments on the perception of voicing in the linguistic context. The results of that experiment will be briefly presented in this paper.

3. The effect of voiceless-voiceless contexts on the perception of voicing in C1C2 sequences

The effect of voiceless-voiceless contexts on the perception of voicing in C1C2 sequences was investigated in synthetic and in natural speech of Dutch. The effect was highly significant (x'2=30.10, df=1, p<.01) difference was found between words containing voiceless-voiceless sequences (C1=(+) C1=(-) C1=(+) C2=(+)) and voiceless-voiceless sequences (C1=(+) C1=(-) C1=(+) C2=(+)).

3.1. A lexical explanation: the listener is inclined to interpret the perceived sounds so that they make an existing word. We may expect, therefore, that the listeners would trend towards the perception of meaningful words, and consequently towards the perception of a voiceless consonant followed by a voiced one, which yields a string of unaltered words. In those cases where a nonword was changed into a word by a shift in the voicing status of one of the members of the C1C2 sequence, we do not expect to prevail. A shift away from the voiceless-voiced responses can thus be expected. Nothing of the kind is observed on the contrary, nonwords show more voiceless-voiced responses than word pairs.

3.2. A phonological explanation: the listener's perception is subject to his knowledge of phonological rules, particularly of the voice alternation rule. We expect that the subjects will perceive more 'regular' sequences in word pairs than in nonword pairs of the same length. This would result in a higher number of voiceless-voiced responses in word pairs with nonvoice-frustrative sequences [6].

The results are to the contrary: we observed more voiceless-voiceless responses in words than in nonwords.

3.3. A phonetic explanation: a change in the sound structure of the context might have affected perception. Since the linguistic and phonological explanations did not adequately predict the observed response patterns, we looked for a third explanation. A shift in the shape of the phonetic context affects the perception of voicing in such sequences. In C1C2 sequences with the same vocal fold setting: vibrating or non-vibrating.

The results showed that in those cases where C1 and/or C2 was changed into a voiceless consonant, the number of voiceless C1's and C2's is increased. A more detailed analysis suggested that changes in the voicing status of C1 were related to changes in C2; responses in the voicing status of C2 to changes in C1. This supports a phonetic explanation. The nonwords were derived from the word pairs by altering the initial (C1) and the final (C2) consonant. Therefore, the only phonetic difference between the word pairs was the C1-C2 word sequence. However, if a difference in the phonetic context affects the perception of voicing in C1C2 sequences, one would expect a difference in the number of responses containing voiced and voiceless consonants.

The results were to the contrary: we observed more voiceless responses than voiceless ones.

4. Discussion

In this section we will discuss four different hypotheses that may explain the results obtained. We will assume that the perceptual mechanism uses its awareness of regularities in speech production for a mechanism like perceptual compensation, and that such a mechanism is sensitive to the perceptual voicing status of C1 and C2. Therefore, we assume that effects of voicing in the context are restricted to the syllable. However, it remains possible that such effects can occur over a larger distance if either of the two hypotheses is applied.
produced with stronger 'voicing' if C₁ (or C₂) is voiceless as compared to the condition where C₁ (or C₂) is voiced. This may be seen as a kind of emphasized articulation of C₁ (or C₂) in the presence of the corresponding voiceless consonant. If, however, the stimulus is ambiguous, as was the case in the present experiment, we may expect that such errors will result in the perception of a more voiceless sequence in a context and a more voice sequence in a context.

A2) Perceptual expectation of coarticular effects

The listener may be inclined to perceive the things he expects, in other words he may be the victim of selective perception. In order to test this result we must assume that the listener expects to hear a voice, e.g., the nonword (the only element to vary in the sentence) has been produced with stronger 'voicing' if C₁ (or C₂) is voiced. This expectation must be based on facts in natural speech. Therefore, we have to assume that voiceless consonants in the context lead to deceiving of (some of) the nearby consonants. From a coarticular viewpoint this is a plausible position.

The hypotheses A₁ and A₂ are mutually exclusive, since they assume opposite facts in production. So, articulatory measurements must enable us to make a choice between the two, or, in case no differences are found, refute both them.

B1) A perceptual-phonological explanation

In this purely perceptual hypothesis we assume that a sequence of speech sounds is recognized in terms of 'phonological packages' to which we may assign the name of 'slices'. The slice effects as the one found in Experiment 1 may occur when a correctly perceived segment (C₁ or C₂) is erroneously attributed to C₀ (or C₁). If this type of erroneous attribution is based on a phonological level, it is likely that the acoustic duration of the interconsonantal phoneme (the vowel) is of no consequence. In this case no effect of (intervening) vowel length is expected. A second factor that can be expected to induce this type of attribution error is the resemblance between the two 'phonological packages'. In the present experiment this resemblance may be caused by the fact that the two phonemes (context and target) have more features in common.

B2) A perceptual-phonetic explanation

In this hypothesis we assume that the error occurs at a more peripheral level: there is a preperceptual integration of the two cues. We claim that the procedure and criteria that we used to assess whether the domain over which phonetic perception is based take place in different terms or in terms of number of phonemes, and thus whether the error occurs in PAS or not, is an indication of the degree of voicing, that may, for example, be operationally defined as the position of C₀ (or C₁) on a 'voicing scale' depending on, e.g., VOT (or VTI). Furthermore, a greater temporal proximity of the context phoneme and the target phoneme (target and context in a sequence) will result in the perception of a more voiceless sequence in a context and a more voice sequence in a context.

The hypotheses B₁ and B₂ make different predictions with respect to the relative duration of the two phonemes and with respect to a possible effect of the degree of voicing in G₁ (or G₂). The production experiment in which the two phonemes are simultaneously present in the context and C₁ (or C₂) leads to more voice in G₁ (or G₂), or on hypothesis A₂ (perceptual expectation) if it leads to less voice in G₁ (or G₂).

In order to decide which of the four hypotheses outlined above are confirmed, we may use measurements from single voiced and voiceless consonants, the vowels intervening the context phoneme and the target phoneme may also promote misattributions if the vowels intervene the context phoneme and the duration of the intervening phoneme (the vowel) is of no consequence. In this case no effect of (intervening) vowel length is expected. A second factor that can be expected to induce this type of attribution error is the resemblance between the two 'phonological packages'. In the present experiment this resemblance may be caused by the fact that the two phonemes (context and target) have more features in common.

The production experiment in which the two phonemes are simultaneously present in the context is the exception experiment has at yet not been carried out.

The production experiment in which the two phonemes are simultaneously present in the context is the exception experiment has at yet not been carried out. The crucial experiment for the choice between hypotheses B₁ and B₂ would be one in which the degree of voice in the context and the duration of the intervening vowels are varied. If perceptual errors result from an erroneous attribution of already recognized phonological features, neither gradations of voicing in the context nor the duration of the vowel phonemes are expected to affect the perception of the Voicing Scale (the degree of voice in C₁ or C₂). On the other hand, the errors are located in the cue integration. In this case, the effects of gradations of voicing and of vowel length are expected to be significant.

The production experiment will be discussed in the next section. The perception experiment has as yet not been carried out.

The production experiment in which the two phonemes are simultaneously present in the context is the exception experiment has at yet not been carried out.

The production experiment in which the two phonemes are simultaneously present in the context is the exception experiment has at yet not been carried out.

The production experiment in which the two phonemes are simultaneously present in the context is the exception experiment has at yet not been carried out. 

Se 92.3.4

References


