PREPLOSIVE FØ IN THE PERCEPTION OF /d/-/t/ IN ENGLISH

K.J. Kohler

Institut für Phonetik, Universität Kiel, Olshausenstr. 40, 2300 Kiel, F.R.G.

ABSTRACT

The importance of level vs. falling FØ contours on prestop vowels for the voiced/voiceless categorization is discussed in the light of perception test data for English "widen/whiten" and compared with corresponding data from German. Over the same set of complementary vowel/stop closure durations, level FØ leads to a greater number of /t/ responses than falling FØ.

INTRODUCTION

Kohler (1982, 1985) presented data from German which support the following points:

1. A measurable FØ contour is related to two factors: a global utterance intonation and local perturbations due to articulatory constraints.
2. In utterance-final disyllabic words of the type "leiden/leiten" ['la^{e}dn/'la^{e}tn] a falling terminal FØ contour changes its global character and consequently its meaning when the FØ peak is located either before/at the initial consonant/vowel boundary or right inside the stressed vowel.
3. In the case of a central peak on the stressed vowel, the FØ fall is delayed by a following voiceless vs. voiced stop consonant.
4. In the case of an early peak on the stressed vowel, the local FØ differences before voiced/voiceless stops disappear.
5. In perception, level vs. level+falling FØ patterns on the stressed vowel favor /t/ and /d/ responses respectively, compared with a continuously falling FØ throughout the stressed vowel.

This paper discusses comparable perception data from English.

PROCEDURE

The sentences "I am telling you I said widen/whiten." ['wa^{e}dn/'wa^{e}tn] with focus stress on the final word were the point of departure for constructing a listening experiment according to the principles outlined in Kohler (1985). Fig.1 represents the speech wave and fundamental frequency of the original sentence "I am telling you I said widen.", which was used for deriving the test stimuli. The duration of [a^e] was reduced from its value of 265 ms in the original "widen" to the value in the original "whiten" by six 10-ms steps (=7 Stimuli). To these vowels closure silences were appended which were increased from 70 ms in six 15-ms steps complementary to the vowel shortening. Three FØ patterns were generated with each vowel duration. (a) Level+falling (119-123-85 Hz); the level section represents the naturally produced fluctuation over the first 100 ms of the original [a^e]; the proportion of level to slope sections stayed the same in all the 7 stimuli. (b) Level (119-123-122 Hz). (c) Linearly falling throughout the vowel (119-85 Hz).

The same ranges of vowel and closure silence durations and very similar FØ patterns (as regards absolute values and timing) were used in the English stimuli as in the German ones.

A group of 12 native Southern British speakers were given the task of classifying the stimulus utterances as "widen" or "whiten" sentences by ticking the appropriate boxes on prepared answer sheets.

RESULTS

Fig. 2 shows the identification functions, as well as the binomial confidence ranges at the 5% level. The response curves for falling and level+falling patterns are very close together, except for the duration ratio of .64, and they are significantly different from level FØ at the low and middle duration ratios: level FØ leads to a higher number of /t/ responses.

DISCUSSION

Basically, the same results as for German have been replicated for English. There are two differences, however:

(a) There are generally more /d/ responses in the English test: the functions are shifted towards shorter ratios.

(b) The curves are closer together, and they are no longer separate for falling and level+falling.

These differences could, of course, be attributed to the different languages, and it might even be objected that such a comparison across languages and test groups is not legitimate. But it is possible to explain the divergencies of the German and English data by reference to Raphael, Dorman and Liberman (1975), who showed that the status of the prevocalic consonant influences the voiced/voiceless perception of postvocalic stops. Their results indicate that the longer the initial voiced formant transitions, the greater the lengthening of perceived vowel duration. In the case of English "widen" vs. German "leiden" the same argument applies since the sequence [w]+[a^e] constitutes a vocalic continuum with extremely long transitions and fuzzy segment boundaries, whereas [l]+[a^e] has a much clearer division. Consequently, [w] increases the perceived vowel duration more than [l]. The general strengthening of [d] responses in the English test is in line with these considerations.

Furthermore, a section of about 40 ms before the segmentation point set between [w] and [a^e] in the original stimulus has a level FØ of 119 Hz. It was not affected in the stimulus construction and therefore stayed the same in all three FØ sets. Thus the linearly falling pattern is preceded by a short level FØ, which, together with the fuzzy segment boundary, prevents it from becoming a different global pattern: linearly falling and level+falling FØ across the segmented [a^e] lead to identical response functions.

In conclusion, we can say that the prosody of the entire stressed syllable, i.e. its total temporal structure as well as its pitch contour, determines segmental voiced/voiceless recognition.

REFERENCES

Kohler, K. J. (1982). "FO in the production of lenis and fortis plosives," Phonetica 39, 199-218.

Kohler, K. J. (1985). "FO in the perception of lenis and fortis plosives," J. Acoust. Soc. Am. 78, 21-32.

Raphael, L. J., Dorman, M. F., and Liberman, A. M. (1975). "The perception of vowel duration in VC and CVC syllables," Stat. Rep., Haskins Laboratories 42/43, 277-284.

ACKNOWLEDGEMENT

I am indebted to Dr Andrew Butcher, Reading, for running the listening test.

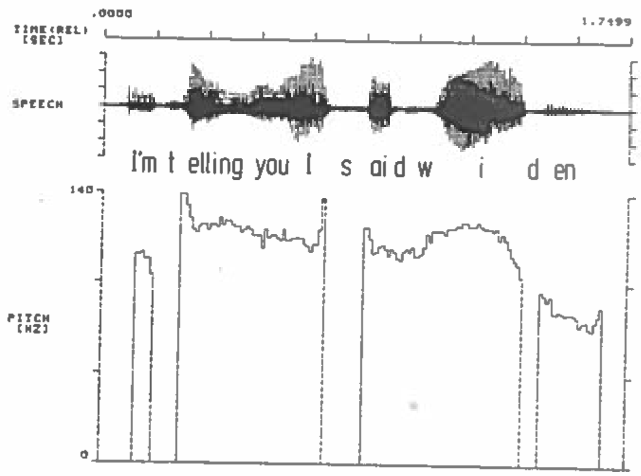

Fig. 1
Speech wave and fundamental frequency of the original sentence "I am telling you I said widen.", which was used for deriving the test stimuli.

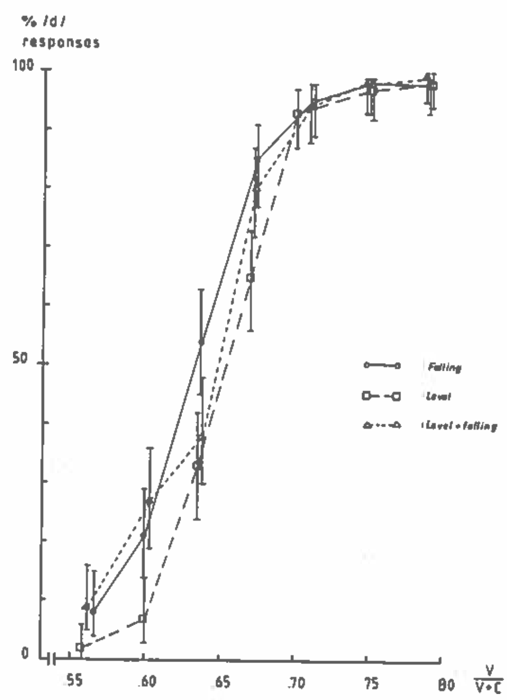

Fig. 2
Percentage /d/ responses as a function of vowel/(vowel + closure) duration ratios for three FØ conditions, and binomial confidence ranges at the 5% level; 12 listeners. At each data point N = 120.