

Wim A. van Dommelen

Institut für Phonetik und digitale Sprachverarbeitung
Kiel, Germany

ABSTRACT

This paper investigates the influence of fundamental frequency (F0) contour on the perception of segment duration in isolated German words. It is shown that the established opinion concerning an increase of perceived segment duration due to a dynamic vs. a flat F0 contour must be modified. The effect appears to be dependent on word structure: While for monosyllables the lengthening effect of dynamic F0 was confirmed, a shortening effect was observed for disyllabic words.

1. INTRODUCTION

Since the experimental work of Lehiste [2] there seems to be a common opinion concerning the influence of the F0 contour in a vocalic segment upon the segment's perceived duration. A dynamic as against a flat F0 contour is generally regarded to lengthen subjective vowel duration. Other investigations [1, 3] showed that the perception of a word-final obstruent as fortis vs. lenis, too, is biased by the character of the F0 contour in the preceding vowel. Here, the increase of lenis judgements due to a moving F0 contour is explained by the subjectively longer vowel duration, biasing in turn phoneme perception towards more lenis.

In previous - hitherto unpublished - experiments I investigated the influence of a varying F0 contour on the perception of vowel quantity in German disyllabic words.

Besides the main cue of vowel duration, phoneme perception was found to be influenced by the F0 pattern of the first syllable's vowel. The direction of the effect, however, was the opposite from what was expected, a moving F0 contour causing a subjective *shortening* of the vowel. Looking for a possible explanation it was noticed that all previous experiments on this issue used either isolated vowels [2] or monosyllables [1, 3]. Two experiments were therefore set up to investigate word structure (monosyllabic vs. disyllabic) effects.

To investigate the influence of segment duration and F0 contour, the German vowel pair /a:/ - /a/ was used, since this is the only vowel quantity opposition which is mainly cued by duration. All the other long/short oppositions are associated with large differences in vowel openness. The monosyllables were represented by the word pair "Aas" (/a:s/, "carrion") - "As" (/as/, "ace"). As a maximally similar word pair containing two syllables "aBen" (/a:sən/, "[we] ate") vs. "Assen" (/asən/, "aces", dative case plural) was chosen. Variation of actual vowel duration should induce listeners to identify the vowel as being phonologically long vs. short. In addition, the vowel's F0 contour was modified (flat vs. flat-falling). According to the findings reported in the literature up to now, a dynamic F0 contour should lengthen subjective vowel duration and

thereby shift the phoneme boundary towards shorter durations. This was taken as a working hypothesis for both monosyllabic and disyllabic words.

Also, to aid interpretation, the duration of the postvocalic fricative was varied. It was expected that the listeners would interpret a shorter fricative as a momentary faster speech rate, implying a subjective lengthening of the preceding vowel. This should apply to mono- as well as disyllabic words.

2. EXPERIMENTAL PROCEDURE

Test 1

A token of the German word "Aas", spoken on a monotone by a trained male speaker, was used for stimulus generation. Prior to electronic splicing, the word was low-pass filtered at 5 kHz and digitized at a sampling rate of 10 kHz. The subsequent manipulations involved the F0 contour and the temporal structure by manipulation of the synthesis frame rate. First, following an LPC analysis two different F0 contours were created. The first one was "flat" (slightly falling from 115 - 110 Hz to avoid an unnatural vowel quality), the second one was initially flat and fell linearly during the second half of the vowel (115 - 113 - 75 Hz). Subsequently, using these two F0 contours the vowel (original duration 285 ms) was synthesized with eight different durations (varying from 110 - 250 ms in 20-ms steps). The fricative was synthesized with its original duration (373 ms) and shortened by approximately one third (250 ms). Each of these two fricatives was spliced with each of the eight tokens of the vowel duration continuum. In total, this test comprised 32 stimuli (8 vowel durations x 2 fricative durations x 2 F0 contours).

Test 2

For the second test, the same speaker from Test 1 produced the

word "aBen" on a monotone. The manipulations of the test word closely followed those from the first test. Two vowel F0 contours were generated, which were identical with those described above. The /ən/ part received a low F0 contour, which fell from 70 to 60 Hz. Its amplitude was attenuated by ca. 13 dB to be auditorily coherent with the falling F0 contour. Furthermore, the duration of the /a:/ vowel (originally 211 ms) was varied from 85 - 190 ms in eight 15-ms steps. As for the monosyllables, the fricative was synthesized with its original duration (211 ms) and a shortened one (141 ms).

For both tests, the stimuli were replicated five times, recorded on analogue tape in a randomized order, and presented to the listeners via a high-quality loudspeaker. The listeners were seated in a sound-treated room and responded by marking one of two alternatives ("Aas"/"As" and "aBen"/"Assen", resp.) on a prepared answer sheet. Twenty-four phonetically naive subjects took part in Test 1 and twenty-two in Test 2. Eight subjects participated in both tests.

3. RESULTS

The results from the first test are presented in Figure 1. It can be seen that the categorization of the vowel as phonologically short vs. long is mainly cued by its physical duration. Second, the effect of F0 contour confirms the finding of Lehiste [2]: A dynamic contour leads to an increase of "long vowel" judgements. Note that this effect holds for both fricative durations. Third, following the expectations, a shorter fricative duration causes an increase of subjective vowel duration.

The results for Test 2 partly parallel those for Test 1. As far as the influence of temporal structure on vowel quantity per-

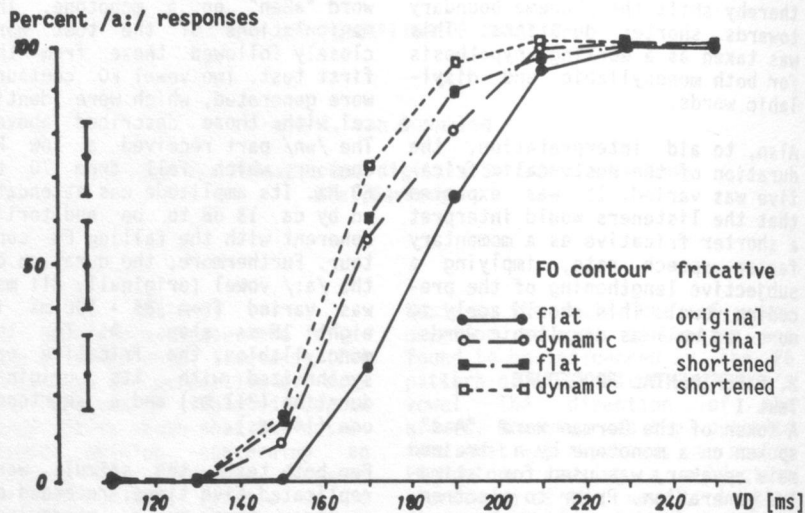


Figure 1
Percentage of /a:/ responses for the monosyllables as a function of vowel duration (VD) for flat and dynamic F0 contours in the vowel and two fricative durations. At each data point $n = 120$. Vertical bars at the left indicate 95% confidence intervals at 25, 50, and 75%.

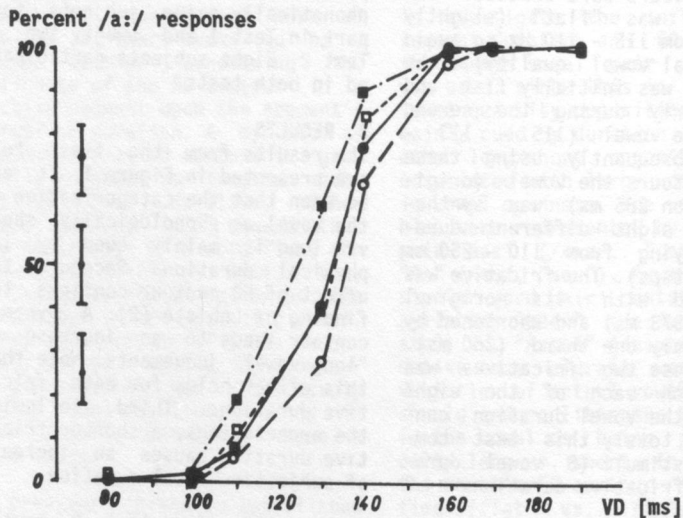


Figure 2
Same as Figure 1, but for the disyllabic words. $n = 110$.

ception is concerned, the results are fully comparable (cf. Figure 2): Whilst phoneme identification primarily depends on actual vowel duration, shortening of the postvocalic fricative biases the location of the phoneme boundary towards shorter durations. In this case, too, perception is additionally influenced by the F0 contour in the vowel. The direction of the effect is, however, exactly the opposite. Here, the presence of a dynamic vs. flat contour causes a decrease, rather than an increase of "aßen" judgements. This occurs consistently for both fricative durations.

4. DISCUSSION

With the results of the test on monosyllables the findings from the literature have been replicated: There was an increase of "long vowel" judgements due to a dynamic vs. flat F0 contour, which certainly reflects the impression of a longer vowel duration in that case. A possible explanation may be the assumption of a production-perception link for voiced segments in absolute final position. This is supported by data from Lyberg [4], who found a positive correlation between segment duration and extension of the F0 fall for a Swedish [dag] syllable in sentence-final position.

The results of the second experiment have clearly falsified the second part of the working hypothesis: The vowel in the disyllabic words was auditorily shortened by a dynamic F0 contour, and not lengthened. Since the experimental conditions (speaker, processing, and test conditions) were identical for both tests, it is highly improbable that the effects are artifactual or due to different perceptual behaviour of the listeners. This is supported by the observation that the effect of varying the fricative duration is identical in both tests.

Further investigations are required to find an adequate explanation for the different interactions of F0 and perceived segment duration in mono- vs. disyllabic words. At the moment, it might be speculated that the crucial feature of the disyllabic word in this respect is the global intonation contour. F0 information in the context surrounding the syllable in question could provide a reference for the listeners. Isolated monosyllables, in contrast, lack this reference. This speculation is supported by the results of preliminary experiments using monosyllables in carrier phrases. It seems therefore that the experiments reported before focused on an exception, rather than the rule.

5. REFERENCES

- [1] GRUENENFELDER, T.M.; PISONI, D.B. (1980), "Fundamental frequency as a cue to postvocalic consonantal voicing: some data from speech perception and production", *Perception and Psychophysics*, 28, 514-520.
- [2] LEHISTE, I. (1976), "Influence of fundamental frequency pattern on the perception of duration", *J. of Phonetics*, 4, 113-117.
- [3] LEHISTE, I.; SHOCKEY, L. (1980), "Labeling, discrimination and repetition of stimuli with level and changing fundamental frequency", *J. of Phonetics*, 8, 469-474.
- [4] LYBERG, B. (1981), "Some consequences of a model for segment duration based on F0-dependence", *J. of Phonetics*, 9, 97-103.