Role of F1 in the Perception of Voice Offset Time as a Cue for Preaspiration.

Jörgen Pind*

*Department of Psychology, University of Iceland, Oddi, IS-101 Reykjavík, Iceland.

Preaspiration, an [h]-like sound inserted between a vowel and a following stop consonant, can be cued by Voice Offset Time (VOffT), a speech cue which is a mirror image of Voice Onset Time (VOT). Previous research has shown that the vowel preceding preaspiration has a much greater effect on the phoneme boundaries for VOffT than does the vowel following VOT. Recent experiments point to the importance of the linguistic context for this effect since preaspiration can only follow a phonemically short vowel in Icelandic. The present paper reports experiments investigating whether this effect could possibly be mediated through the influence of F1 on perceived voicing.

INTRODUCTION

Preaspiration, an [h]-like sound inserted between a vowel and the following stop closure, is a rare phonetic contrast in the languages of the world. Previous research has shown that the perception of preaspiration in Icelandic is in many ways similar to the perception of Voice Onset Time, VOT. Thus Pind (1) showed that preaspiration can readily be cued by Voice Offset Time (VOffT) a speech cue which is a mirror image of VOT. In one respect, however, the perception of VOffT was shown to diverge from the perception of VOT in that VOffT is much more sensitive to the duration of the preceding vowel than is VOT to the duration of the following vowel. Thus lengthening the duration of the vowel from 132 to 232 ms, in a /ba-pa/ continuum, moved the phoneme boundaries for VOT from 29.7 to 35.5 ms. A comparable lengthening of the vowel preceding VOffT in an /ap-ahp/ continuum moved the phoneme boundaries for VOffT from 30.7 to 52.6 ms of VOffT (2).

This discrepancy in the perception of VOT and VOffT could plausibly be explained in various ways (3). Thus it is possible that this effect might reflect the operation of psychoacoustic factors such that the vowel preceding preaspiration would to some extent serve as a masker of the weaker aspiration, a condition which would presumably not hold for the perception of VOT (4). Recent research, using non-speech analogs of VOT and VOffT, do not lend support to this hypothesis but rather highlight the importance of the linguistic context of VOT and VOffT in Icelandic (3).

Preaspiration, as already mentioned, follows the vowel in Icelandic. Icelandic distinguishes short and long vowels and consonants. These show complementary durations, thus a long vowel is followed by a short consonant, a short vowel by a long consonant or a consonant cluster. Preaspiration along with the following closure forms one such cluster, thus the phonemic structure of Icelandic words is such that preaspiration always follows a phonemically short vowel. Such a situation does not hold for VOT. VOT, coming at the beginning of the syllable, can either precede a phonemically short or a phonemically long vowel.

The importance of vowel quantity on the location of the perceptual boundaries for VOffT were brought out in a recent study (3) where it was shown that the boundaries for preaspiration where shifted to longer values of VOffT regardless of whether the vowel quantity was increased by increasing the vowel duration or by changing the spectrum of the vowel towards that of a long vowel while keeping the duration constant. This provided strong evidence for the hypothesis that the increased VOffT needed to cue preaspiration with lengthened vowels is in fact due to changes in the perceived quantity of the vowel.

A possible confound was noted in this experiment, namely the lower F1 associated with the phonemically long vowel [ɛː], compared to the short phoneme. It is known that the perception of VOT is influenced by the value of F1 at voicing onset with a lower F1 needing a longer VOT at boundary (5). Thus it could be that the lower F1 in the phonemically long [ɛ] was contributing to the longer VOffT boundaries. The present study addresses this issue.

METHOD

The stimuli for this experiment were, for the most part, modeled on those of Kluender (5) and synthesized with the SenSyn synthesizer, a version of the Klatt synthesizer (6). There were four VOT /ga-ka/ continua and four VOffT /ak-ahk/ continua, distinguished by the frequency of F1 at voicing onset (or offset in the case of VOffT), and the duration of the F1 transitions. There were four series in all: A: F1 onset/offset of 450 Hz, transition duration of 20 ms; B: F1 = 150 Hz at onset/offset, transition duration = 40 ms; C: F1 at onset/offset = 450 Hz, transition duration = 40 ms; D: F1 = 150 Hz at onset/offset, transition duration 60 ms. Steady state frequencies of the vowel [a] were fixed at 750, 1275 and 2425 Hz respectively for F1, F2 and F3. The vowel was synthesized with a breathy quality (parameter AH = 40 dB). VOT and VOffT were synthesized with a combination of aspiration and frication, the latter in the frequency region of F2 and F3, while simultaneously cutting off the amplitude of voicing and increasing the bandwidth of F1 from 60 to 200 Hz. The transitions for F2 and F3 were 40 ms long and consisted, in the case of VOT, of a linear fall of F2 from 1900 to 1275 Hz, while F3 rose linearly from 2100 to 2425 Hz. For VOffT the transitions were mirror-images of those for VOT. Overall duration of the VOT stimuli was 132 ms. 120 ms of silence and a velar burst were added to the 132 ms long vowels in the case of the VOffT stimuli. Twelve subjects participated in the listening test where they were to say, in the case of VOT, whether the syllable was /ga/ or /ka/ and, in the case of VOffT, whether the syllable was /ak/ or /ahk/.

RESULTS AND DISCUSSION

Figure 1 shows pooled identification curves for all participants in both experiments.

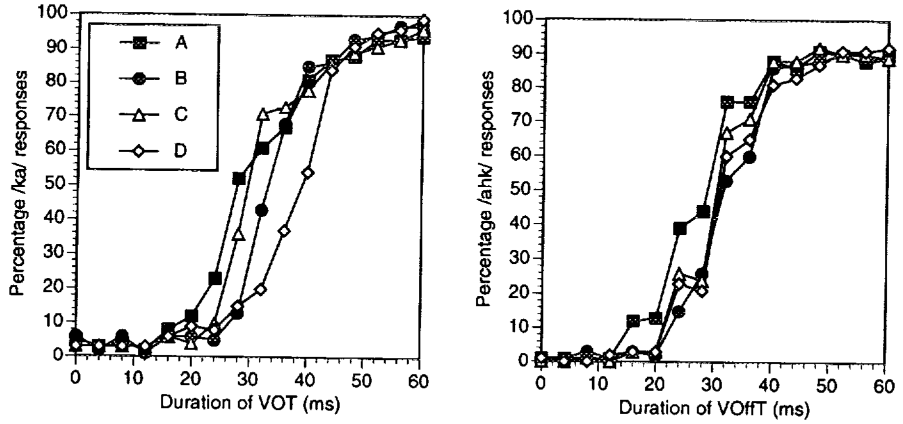

FIGURE 1. Pooled identification curves, VOT on the left, VOffT on the right. Figure legend denotes stimulus series.

Average phoneme boundaries range from 31.3 ms (Series A) to 37.6 ms (Series D) in the VOT condition. Differences are in the same direction as in Kluender (5) though of less extent in the present experiment, which is to be expected given that in Kluender's D-series the transition was 80 ms long. (Transitions of this duration sound highly unnatural in VOffT stimuli and were therefore not used.) Clearly, the lower the F1 at onset the more VOT is needed to cue aspiration. In the VOffT condition the results are similarly patterned, though the effect of F1 at offset is somewhat less with average phoneme boundaries at 27.6 ms in Series A and at 31.2 ms in Series D. The difference here amounts to 3.6 ms compared to 6.3 ms in the case of VOT, though this difference is not statistically significant, $t(11)= 1.24$, $p = 0.24$. In any case it is clear that F1 frequency at voice offset can not explain the noticeable effect of vowel duration on the perception of preaspiration previously shown for Icelandic (2–3).

ACKNOWLEDGEMENTS

This research was supported by the Icelandic Science Foundation and by the Research Foundation of the University of Iceland. I am grateful to Thórunn H. Halldórsdóttir for help in running the listening tests.

REFERENCES

1. Pind, J. *Acta Psychologica* **89**, 53–81 (1995)
2. Pind, J. *Quarterly Journal of Experimental Psychology* **49A**, 745–764 (1996).
3. Pind, J. *Journal of the Acoustical Society of America*, in press.
4. Bladon, A. (1986). Phonetics for hearers. G. McGregor (ed), *Language for hearers*. Oxford: Pergamon Press, 1986, pp. 1–24.
5. Kluender, K. R. (1991). *Journal of the Acoustical Society of America*, **90**, 83–96.
6. Klatt, D. H., & Klatt, L. C. (1990). *Journal of the Acoustical Society of America*, **87**, 820–857.