Preaspiration and Perceived Vowel Duration in Norwegian

Jacques Koreman¹, William J. Barry² and Marte Kristine Lindseth¹
¹Department of Language and Communication Studies, NTNU, Trondheim
²Institute of Phonetics, Saarland University, Saarbrücken

Abstract

This article presents an experiment to investigate the perceived duration of Norwegian vowels before [d] versus preaspirated [t]. It is shown that preaspiration contributes to the perceived duration of the vowel before [t]. The general implications of this finding for phone segmentation and for phonetic research using vowel duration measurements are discussed.

Introduction

In a production study of American English, Peterson and Lehiste (1960) found that vowel duration including aspiration after phonologically voiceless or fortis plosives (308 ms) is longer than after phonologically voiced or lenis word-initial plosives (274 ms); if aspiration is excluded, it is shorter (251 ms) on average in a set of 68 minimal pairs. In a perception study for German (results as yet unpublished), a corresponding effect on perceived duration was shown: the vowel duration after a lenis plosive was judged equal to the duration of the vowel plus half of the aspiration after a fortis plosive.

Comparable to the production data above, the vowel incl. preaspiration before a tense plosive is longer than the vowel before a lax plosive in “preaspirating” dialects of Norwegian (Van Dommelen, 1999), while the vowel excluding preaspiration is shorter; but there is substantial variation across dialects. In perception, the perceptual effect of preaspiration on phonological categorization has been investigated by Moxness (1997), who found that the perceived (phonological) vowel length is not affected by the presence or absence of preaspiration in V:C versus VC: stimuli. Van Dommelen (1998) showed an effect of preaspiration on the perception of plosives as fortis versus lenis in Norwegian.

The present study investigates the perceived phonetic vowel duration in stimuli containing preaspirated versus fully voiced vowels. We hypothesize that, similar to aspiration, preaspiration influences the perceived vowel duration. More specifically, our goal is to evaluate how much of the pre-aspiration is perceived as part of the vowel.

Method

We shall first discuss the selection of the stimulus pairs and then describe their preparation for the perception experiment, followed by a description of the experiment itself.

Selection of the stimuli

Ten repetitions of two sets of /CV:Cα/ and /CVC:α/ stimuli were recorded with all combinations of /i,a,u/ followed by /p,b; t,d; k,g/ (and with the same initial C in each minimal pair differing in [voice] for the second consonant). The stimuli were presented for reading in random order on a computer screen using a PowerPoint presentation, and were recorded directly onto hard-disk in a studio, with a sampling frequency of 44 kHz and a 16-bit amplitude resolution. Comparison of the stimuli showed that preaspiration after short vowels is generally longer than after long vowels. The vowel /α/ showed no supraglottal friction, which did sometimes occur with close vowels, especially /i/. To maximize the presence of true preaspiration, we selected “batte”-“badde” from the list as the single stimulus pair for our perception experiment. An additional consideration was that these are both nonsense words, so that the listeners are not affected by familiarity or frequency of the stimuli.

Figure 1. Segmentation of vowel in “batte” into modally voiced, breathy and preaspirated portions

To investigate how perceived vowel duration is affected by preaspiration, we used ten repetitions of the stimuli spoken by a single, male speaker from Stavanger. Figure 1 shows an example of the segmentation of the vowel in “batte” into modally voiced, breathy voiced and...
preaspirated portions, for which praat was used. Since there was typically a breathy voiced signal portion in the transition from modal voice to preaspiration, this was included as a separate factor in the perception experiment.

**Preparation of the stimuli**

Two native Norwegian listeners (both MA students of Phonetics) could not distinguish the signal portions of the two sets of stimulus words from the release to the end of the word. Thus, “batte” and “badde” are differentiated by the stressed vowel and the following closure.

The vowel of the nonsense words of the “batte” type all consisted of modal voicing, followed by breathy voicing and preaspiration. The vowel in “badde” consisted entirely of modal voicing which continued into the closure of the following /d/. The stimuli for the listening test were adapted so that the durations of the vocalic portions were carefully controlled:

**Stimulus 0**: 
- The duration of the vowel in “badde” equals that of the modally voiced + breathy voiced + preaspirated portions of the vowel in “batte”.

All stimulus pairs were selected such (from the ten repetitions) that the two vowels in a pair were similar. If this was not possible, the two stimuli were changed by deleting or adding single glottal periods until the vocalic portions of interest had more or less the same durations.

The following stimuli have increasingly longer vowels (from vowel onset until the closure of the following consonant) in “batte” than in “badde”:

**Stimulus 1**: the duration of the vowel in “badde” equals that of the modally voiced + breathy voiced + half of the preaspirated portion in “batte”.

**Stimulus 2**: the duration of the vowel in “badde” equals that of the modally voiced + breathy voiced portions in “batte”.

**Stimulus 3**: the duration of the vowel in “badde” equals that of the modally voiced portion in “batte”.

Notice that in the last three stimuli, the “batte” stimulus word will be longer than “badde”, namely by the other half of the preaspirated portion (which is still present!) in stimulus 1, the whole preaspirated portion in stimulus 2, and the breathy voiced + preaspirated portions in stimulus 3.

In addition to the above three conditions there are two conditions in which the vowel in “batte” is relatively shorter:

**Stimulus -1**: the duration of the vowel in “badde” is longer than the total vowel duration in “batte” by the duration of the breathy voiced portion in “batte” (but the vowel is modally voiced throughout).

**Stimulus -2**: the duration of the vowel in “badde” is longer than the total vowel duration in “batte” by the duration of the breathy voiced + half of the preaspirated portion in “batte” (but the vowel is modally voiced throughout).

We did not include a condition stimulus -3 (same as stimulus -2, but with an even longer vowel in “badde”) because the vowel in that case was always perceived as longer than the total vowel in “batte” in preliminary listening. Inclusion would have increased the number of stimuli, without giving additional information.

Two stimulus sets were prepared: in set A, a stimulus pair was selected for each condition from the ten different realizations of “batte” and “badde” which fulfilled its vowel duration criteria (cf. Figure 2). The advantage of this stimulus set is that the stimuli were produced naturally, i.e. the (majority of the) stimuli did not need to be manipulated (by dropping or copying glottal periods) to obtain the vowel durations according to the scheme in Figure 2. The disadvantage of those stimuli is that in addition to the durational differences there may be other factors which influence the perception of vowel duration.

For this reason, we also created a set B, in which only one stimulus pair was selected as a basis for the perception experiment. The pair was chosen so as to fulfill condition 0, i.e. the vowel in “badde” had the same duration as the total vowel duration in “batte”. To derive the
other conditions, the stimuli were manipulated by inserting or deleting glottal periods from the modally voiced portion of the vowel. The “batte” and “badde” stimuli were manipulated equally strongly (e.g. two glottal periods inserted into “badde” and two dropped from “batte” for condition -1).

**Perception experiment**

Eight listeners listened to stimulus set A. The same listeners, plus another four, listened to stimulus set B. The listeners were all native Norwegians between 25 and 60 years old, and had no reported hearing problems.

The listeners’ task was to judge which of the two stimuli in each pair was longest and respond by checking the corresponding box on a response form (on paper) to indicate their perception. There was no “equal duration” choice, since we wanted to prevent the listeners from using this option too often. The listeners could hear the stimulus pair over headphones as often as they liked by moving the mouse over a loudspeaker symbol in the PowerPoint presentation. “Mouse Over” was used instead of “Mouse Click” to prevent the disturbing sound of mouse clicks.

The “batte”-“badde” pairs were offered in both orders, balanced across ten lists (repetitions). Within each list, the stimulus pairs were offered 10 times in different pseudorandomized order. The pseudorandomization consisted in ensuring that two consecutive stimuli were more than 2 apart, i.e. Stimulus 1 could not be followed by Stimulus 0 or 2. The total number of stimuli was 6 conditions × 2 orders × 10 repetitions = 120 stimuli. Each list was preceded by five and followed by three filler items.

**Results**

The two versions of the experiment using stimulus sets A and B led to substantially different results (see Figure 3). In general, the upward trend in “longer vowel” responses from Stimulus -2 to Stimulus 3 corroborates our hypothesis. But for set A, the trend is weak and the number of responses where the vowel in “batte” is considered longer than that in “badde” never exceeds 50%. That is, the vowels in the “badde” stimuli are always judged longer than those in “batte”.

For set B, where the stimuli were all derived from a single pair of stimuli, the trend is much clearer, and shows a clear transition from 3% “longer vowel” responses for Stimulus -2 to almost 94% for Stimulus 3.

**Differences between stimulus conditions**

Separate analyses of variance for the two stimulus sets showed that order of the stimuli within the pair had no significant effect on the perceived relative vowel duration in the two stimuli, nor did order interact with stimulus condition. The difference between the six stimulus conditions, however, was highly significant for both stimulus sets (set A: $F(0.001;5,84)=10.21$; set B: $F(0.001;5,132)=146.83$).

For set A, Scheffé’s post-hoc tests resulted in three homogeneous subgroups (-1,-2,0 < 0.3 < 3.2), where the middle subgroup is almost significant at 5%. The tendency is therefore that the perceived vowel duration in stimulus conditions 2 and 3 differs from the other conditions. But the vowel in “batte” is mostly perceived as shorter than that in “badde”, as noted at the beginning of this section.

For stimulus set B, there is also a division into three homogeneous subgroups (-2,-1,0 < 0.1 < 2.3), but the effect is more consistent with our hypothesis, with the percentage of “longer vowel in batte” responses increasing from stimulus condition -2 (3%) to stimulus condition 3 (94%). A sudden change in the response is visible going from condition 1 to condition 2.

**Differences between the listeners**

Clear differences can be observed between the listeners. In the responses to stimulus set A (see Figure 4), listener PL for instance follows the expected pattern with a steady increase in the number of “longer vowel in batte” responses from stimulus condition -2 to condition 3, and
this subject had more than 50% such responses in condition 3. Listener SH on the other hand does not seem to be influenced by the differences between the stimuli, with around 50% “longer vowel in batte” responses for all stimulus conditions.

Figure 4. Individual listener response percentages (of “longer vowel in batte”) for set A

This shows that it was not only the stimuli which created the differences, although on the other hand all subjects behaved roughly the same for stimulus set B (see Figure 5), which was more strongly controlled in that all stimuli were derived from a single stimulus pair.

Figure 5. Individual listener response percentages (of “longer vowel in batte”) for set B

Discussion

The two stimulus sets show very different results. Set B seems to be most reliable in the observed tendencies, both across stimulus conditions and across subjects. These results show that the vowel in “batte” is judged longer than that in “badde” when the duration of the modally voiced + breathy voiced + half of the preaspirated portion of the vowel together exceed the duration of the vowel in “badde”. In other words, the whole breathy voiced and half of the preaspirated portion of the vowel contribute to its perceived length. This corresponds to previous observations in concerning (post-)aspiration.

But these observations are not corroborated by the results for set A. In this stimulus set, the variation across stimulus conditions is much smaller and the behaviour of individual listeners shows more variation. This may indicate that the listeners rely on different cues for vowel duration, which may also differ across the stimulus pairs – remember that the stimuli in the listening experiment were not all based on the same pair as in set B. Also, there was a strong preference for “longer vowel in badde” responses in all conditions for set B. We were not able to find a reason for this, despite close inspection of the stimuli.

The results highlight a possible inconsistency in segmentation conventions, since aspiration (also used in Norwegian after fortis plosives) is normally considered part of the preceding plosive which causes it, whereas preaspiration is normally segmented as part of the vowel instead of the following plosive. The results for their effect on perceived vowel duration which are reported here and the unpublished results for aspiration show, however, that both affect the perceived duration of the vowel. Using traditional segmentation criteria can therefore lead to wrong conclusions if the segmentation in phonetic studies on vowel duration is used to make inferences about the effect of vowel duration in perception.

Of course, this study was limited in its approach. Other consonantal places of articulation and the speakers’ sex, which have been shown to differ in production studies (e.g. Helgason and Ringen, 2008), should be taken into consideration in perceptual studies.

References


