PERCEIVED SWEDISH VOWEL QUANTITY: EFFECTS OF POSTVOCALIC CONSONANT DURATION

Dawn M. Behne†, *Peter E. Czigler*‡ *and Kirk P. H. Sullivan*‡

†Norwegian University of Science and Technology, N-7043 Trondheim, Norway
‡Umeå University, S-901 87 Umeå, Sweden

ABSTRACT

In the production of Swedish, vowel quantity is known to be realized in the vowel, but also affects duration of a postvocalic consonant. The goal of this study is to examine the use of postvocalic consonant duration as a perceptual cue to vowel quantity. Listeners' responses and reaction times were recorded for synthesized materials in which the vowel spectra and duration were kept constant and the postvocalic consonant duration was adjusted. Results show no indication that listeners actively used the duration of a postvocalic consonant to identify vowel quantity. These findings suggest that adjustments in postvocalic consonant duration in Swedish productions may be temporal artifacts of the preceding vowel quantity rather than reflecting linguistically relevant information.

1. SWEDISH VOWEL QUANTITY

The vowel system of Swedish has traditionally been described as having contrastive vowel quantities, with "short" (e.g., [a] in [tak] *tack* "thanks") and "long" (e.g., [å:] in [t å:k] *tak* "roof") vowels characterized as similiar in quality but phonologically distinguished by their relative lengths (e.g., [1]).

Phonotactically in Swedish, a short vowel quantity will only occur in a closed syllable (e.g., [a] in [tak:] *tack* "thanks"), whereas a long vowel quantity can occur in either an open or closed syllable (e.g., [å:] in [tå:] *ta* "take", or in [tå:k] *tak* "roof") (e.g., [1]). This pattern is highly consistant in Swedish.

In speech production, the distinction between short and long vowel quantities is realized in the acoustic signal corresponding to both the vowel and, if present, a postvocalic consonant. The duration of a long vowel quantity extends over more time than a short vowel quantity which, in turn, may allow more time for an articulation using greater extremes of the vocal space, and consequently may also affect the vowel spectrum. In particular, long vowel quantities are generally known to be articulated with more closure than short vowels, with the open articulation of [Å:] and [a] being a possible exception.

Swedish vowel quantity is also known to affect a postvocalic consonant in speech production. The duration of a postvocalic consonant has been regularly observed [1] to have an inverse relationship to the vowel duration, so that following a short vowel the duration of a postvocalic consonant tends to be relatively long (e.g., [k:] in [tak:] *tack* "thanks"), whereas following a long vowel the postvocalic consonant duration tends to be shorter (e.g., [k] in [tå:k] *tak* "roof").

In a classic study on Swedish, Hadding-Koch and Abramson [2] investigated whether vowel duration or spectral attributes of a vowel had the more dominant perceptual role in distinguishing short and long vowel quantities. For three vowel pairs, tape recordings were carefully spliced with differences of 10-15 ms, resulting in approximately 5-8 steps from the long to the short vowel quantitiy. Although the role of spectral characteristics could not be excluded from being an important perceptual cue, their results show that length is the primary parameter distinguishing Swedish vowels.

Recent studies[3][4] reexamined the effects of vowel duration and spectrum on perceived vowel quantity identification with the goal of identifying thr role of F1 and F2 for identifying vowel quanity. Based on natural Swedish productions, three sets of 100 /kVd/ materials were synthesized, one set for each of three "long-short" vowel pairs [4]. Within each set 10 stepwise adjustments of vowel duration and 10 stepwise adjustments of the first two vowel formants (F1 and F2) were made from the long vowel quantity to the short one, resulting in a total of 100 synthesized /kVd/ items for each of the three vowel pairs. The closure duration of the postvocalic /d/ was kept constant. Subjects' identification are presented again here in Figure 1. The results illustrate that Swedish listeners use vowel duration more than spectral information to identify the quantity of the non-low vowel pairs (/i:/-/I/ and /o:/-/o/), whereas for low vowel pair (/a:/-/a/), they use both vowel duration and spectral attributes of the vowel.

2. CURRENT STUDY

Previous research has shown that Swedish speakers and Swedish listeners both use vowel duration, and in some cases the spectrum of a vowel, to distinguish "long" and "short" vowel quantities.

Swedish speakers also regularly adjust the duration of a consonant following long and short vowel quantities [1]. However, to date, the perceptual role of postvocalic consonant duration as a cue for distinguishing long and short vowel quantities has not been examined. Since long vowel quantities can occur in an open syllable, the use of postvocalic consonant duration as a perceptual cue to vowel quantity is questioned.

The goal of the present study is to examine the use of postvocalic consonant duration as a perceptual cue to vowel quantity in Swedish. The same methodology as [4], is used so that responses from the current study can be directly compared with those of the earlier experiment.

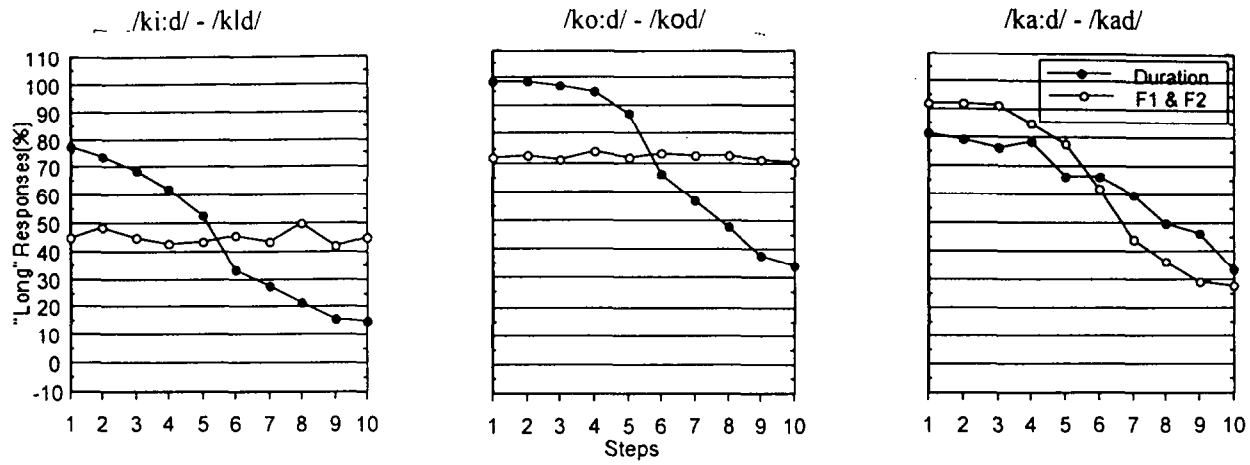

Figure 1: Percent "long" responses from [4] are plotted for 10 duration (-•-) steps and 10 spectral (-o-) steps. Step 1 corresponds to the "long" vowel quantity and step 10 corresponds to the "short" vowel quantity for /ki:d/-/kld/, /ko:d/-/kod/, and /ka:d/-/kad/.

Set	Vowel				Closure duration of /d/		
	Duration	F1	F2	F3	Start (step 1)	Step size	End (step 10)
/ki:d/ - /kld/	121ms	265 Hz	2224 Hz	3224 Hz	60 ms	7	122 ms
/ko:d/ - /kod/	145 ms	317 Hz	666 Hz	2446 Hz	56 ms	8	128 ms
/ka:d/ - /kad/	149 ms	579 Hz	1139 Hz	2288 Hz	54ms	5	102 ms

Table 1. Parameter settings of the vowel and postvocalic consonants for the three sets of resynthesized materials. In each set the vowel duration and F1-F3 were held constant as the closure duration of postvocalic /d/ was adjusted in 10 steps.

2. METHOD

2.1 Materials

The recordings and measurements used here are the same those used in [4].

Recordings

The set of six /kVd/ words from [4] were used as targets. Each word was phonotactically possible in Swedish and contained one of the vowels /I, o, a, i:, o:, a:/.

A young adult native male speaker of the Stockholm dialect of Swedish was recorded producing 10 random repetitions of the six target words in the sentence "Jag sa ___ igen." ("I said ___ again.") at his natural speaking rate.

Measurements

From the 10 productions of each target word, ESPS/*waves+*™ was used to measure the vowel duration, the first three formant frequencies of the vowel (F1, F2 and F3) measured at the center of the vowel's most evident steady state, and the closure duration of the postvocalic /d/.

For each repetition of the six target words, the mean value of these measures was calculated and the production which best corresponded to the mean values was chosen to be used as the basis for resynthesis. These most representative items will be referred to as "selected productions"

Synthesis

Using the Kay Elemetrics LPC Parameter Manipulation/ Synthesis program, the selected productions of the six target words and their measured values were the basis for resynthesizing three series of ten words (/ki:d/ to /kld/, /ko:d/ to /kod/, and /ka:d/ to /kad/).

Starting from the selected production of /ki:d/, /ko:d/ and /ka:d/, the closure duration of the postvocalic consonant was adjusted in equal-sized steps toward the measured postvocalic closure duration of the selected productions of /kld/, kod/ and /kad/ respectively. The closure durations and step sizes are summarized in Table 1.

In each series the vowel duration and the first three formant frequencies of the vowel were held constant by basing the synthesis on the averaged vowel measurements between the two selected productions used for the series. The vowel durations and first three formant frequencies used are summarized in Table 1.

2.2 Procedure

Twenty native speakers of Swedish between 20 and 38 years old participated in the study.

Subjects were seated wearing headphones at a computer terminal with a monitor and mouse. For each trial, subjects heard a synthesized word and at the same time two real words (vid - vidd, nåd - nådd, or vad - vadd) were presented on the monitor. The two words on the monitor differed in vowel quantity and had the same vowel quality and postvocalic consonant as the target words which the synthesized items in that series were based on.

Subjects were instructed to use the mouse to click on the visually presented word which rhymed with the one they heard. They were asked to respond as quickly as possible and were allowed up to 10 seconds to respond before the beginning of the next trial, although subjects rarely encountered this upper limit.

Subjects heard 5 randomized repetitions of each synthesized word, a total of 150 items (3 sets x 10 items x 5 repetitions). Before starting the experiment, subects had three practice trials, and after each set of 50 trials, subjects had the opportunity to take a short break.

Subjects' responses and reaction times for each trial were analyzed to determine the extent to which the closure duration of the postvocalic consonant affects the perceived quantity of the preceding vowel.

3. RESULTS

The mean percent responses that were "vid", "nåd", or "vat" was calculated for each condition. These are referred to as "long responses" in the following discussion. The percent long responses and reaction times for each of the ten consonant steps in each series are shown in Figure 2. Standard deviations is shown by the vertical bars.

Do listeners used the closure duration of the postvocalic consonant to identify the quantity of the preceding vowel?

The effect of postvocalic closure duration on the percent of long responses for each of the three series-is shown by the solid lines in the top row of Figure 2. No clear categorization boundary between long and short vowel quantities could be identified based on postvocalic closure duration, evidenced by the relatively flat horizontal curves In none of the three series was a reliable difference in percent long responses observed (/ki:d/-/kld/: [F(9,200)=0.22, p=0.991], /ko:d/-/kOd/: F(9,200)=0.463, p=0.8981], /kÅ:d/-/kad/: [F(9,200)=0.049, p=1.000]).

This finding is supported by the reaction time results presented in the bottom row of Figure 2. Although there is some variation in reaction times across the 10 steps of the three series, they are not systematic, suggesting that the cognitive effort involved in the task was relatively consistant across the 10 steps. No reliable difference was observed [F(9,200) = 0.544, p = 0.841], /ko:d/-/kOd/: [F(9,200) = 0.247, p = 0.987], /kÅ:d/-/kad/: [F(9,200) = 0.437, p = 0.914]). Reaction times were also relatively long (around 1300 ms) suggesting that the identification of vowel quantity based only on postvocalic consonant duration was a relatively difficult task.

The individual listeners responses also support that the postvocalic consonant duration is a strong acoustic perceptual cue for identifying vowel quantity. With the possible exception of one subject whose percent long responses hint (but not very clearly) at a category boundary, the subjects' responses are either essentially the same across the 10 consonant duration steps, or completely irregular.

These findings suggest that listener's generally do not use the duration of a postvocalic consonant as a cue to vowel quantity.

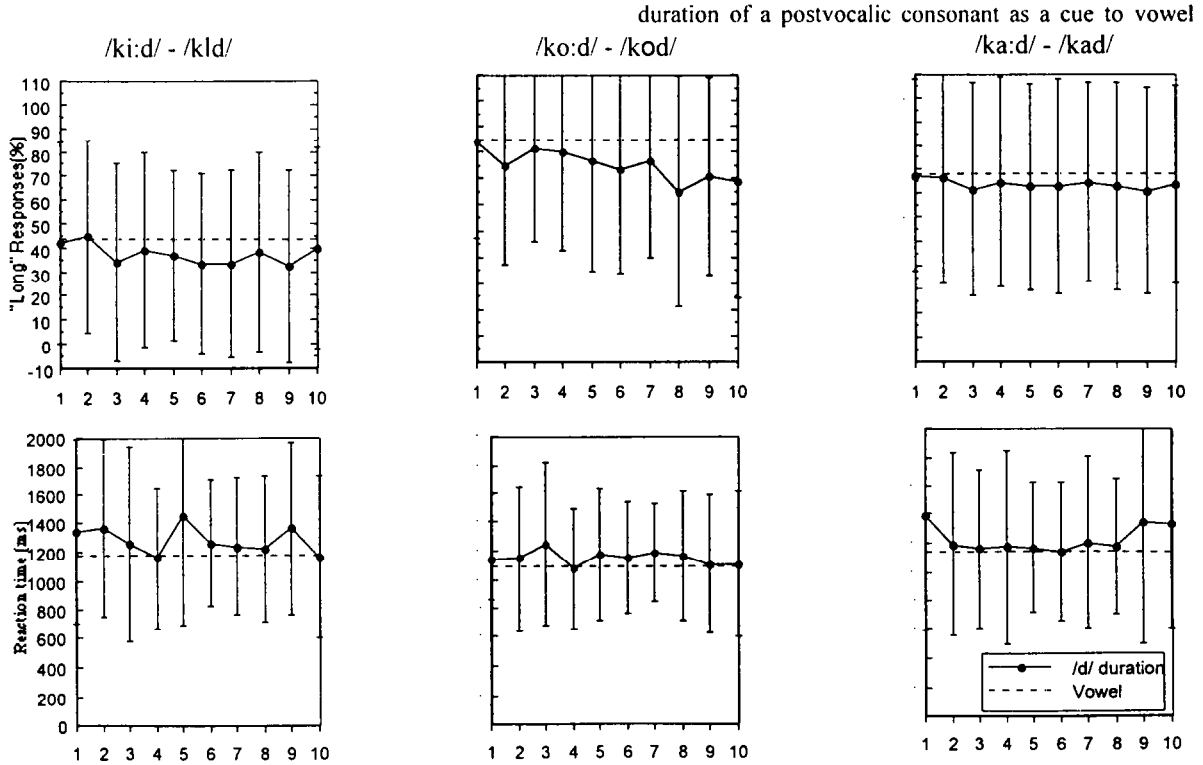

Figure 2: Percent "long" responses and reaction times are plotted for the 10 steps of postvocalic /d/ closure duration for /ki:d/-/kld/, /ko:d/-/kod/, and /ka:d/-/kad/. The correspondingresults in [4] are presented for comparison.

Are there differences across vowel qualities?

Even though the percent long responses do not differ systematically across the 10 steps for any one of the three series, the mean percent long responses does differ among the three vowel quality series [F(2,627) = 32.110, p = 0.0001]. This difference is likely due to the nature of the preceding vowel. Recall that in the current study, the duration and F1- F2 frequencies of the vowel within each series was adjusted to intermediate the long and short vowel quantities. This corresponds to reponses from [4] which were intermediate to steps 5 and 6 of the vowel duration and vowel spectra. As the results above in Figure 1 illustrate, the mean percent long responses at steps 5 and 6 of vowel duration and F1-F2 in [4] were not at 50% (μ for /ki:d/-/kId/=43.8%; μ for /ko:d/-/kod/= 74.9; μ for /ka:d/-/kad/=68.2). These differences are discussed in [4], but also establish the mean percent long responses that expected in the current study if postvocalic consonant duration has no effect on perceived vowel quantity. These means are plotted as flat dotted lines Figure 2, and as can be seen, they appear to be comparable to the results from adjusting the postvocalic consonant duration in the current study.

To test this, confidence intervals were calculated for the 10 steps of postvocalic consonant duration for each of the three vowel quality series. The mean percent long responses and reaction times between steps 5 and 6 in [4] were compared with the confidence intervals and are presented in Figure 3. Results show that the listeners' mean responses and reaction times between steps 5 and 6 in [4] consistantly fall within the confidence intervals at the 10 steps of adjusted postvocalic consonant duration in each of the vowel three vowel series.

4. CONCLUSIONS

Swedish speakers have regularly been observed adjusting the duration of a postvocalic consonant to be relatively short following a long vowel quantity and to be relatively long following a short vowel quantity. This inverse relationship between the vowel and postvocalic consonant could, in principle, offer Swedish listeners a perceptual cue to the vowel's quantity.

The current results from adjustments of stop closure durations strongly suggest that listeners do not use the duration of postvocalic consonants to identify vowel quantity in Swedish, even in the absense of durational or spectral information from the vowel, indicating that the inverse durational relationship between a vowel and postvocalic consonant in Swedish productions does not have an immediate linguistic role as an acoustic cue to vowel quantity. Notably, the relatively inverse durational relationship between a vowel and postvocalic consonant is not unique to Swedish, or languages with distinctive vowel quantity (e.g., [4]) suggesting that adjustments of postvocalic consonant duration in Swedish productions may be temporal artifacts of the preceding vowel quantity rather than an acoustic cue for linguistically relevant information.

5. ACKNOWLEDGEMENTS

The authors thank Ola Andersson and Thierry Deschamps from Umeå University for their technical assistance.

6. REFERENCES

1. Elert, C-C., *Phonologic studies of quantity in Swedish*, Stockholm: Almqvist & Wiksell, 1964.

2. Hadding-Koch, K., and Abramson, A., "Duration versus spectrum in Swedish vowels: some perceptual experiments", *Studia Linguistica* 2: 94-107, 1964.

3. Behne, D. M., Czigler, P. E, and Sullivan, K. P. H., "Acoustic characteristics of perceived quantity and quality in Swedish vowels", *Speech Science and Technology '96*, Adelaide, Australia: 49-54, 1996.

4. Behne, D. M., Czigler, P. E, and Sullivan, K. P. H., "Perceived vowel quantity in Swedish: Effets of postvocalic voicing", *Proceedings of the 16th International Congress of Acoustics and the 135th Meeting of the Acoustical Society of America*; 2963-64, 1998.

5. House A., and Fairbanks G. "The influence of consonant environment upon the secondary acoustical characteristics of vowels." *Journal of the Acoustical Society of America* 25: 105-113, 1953.

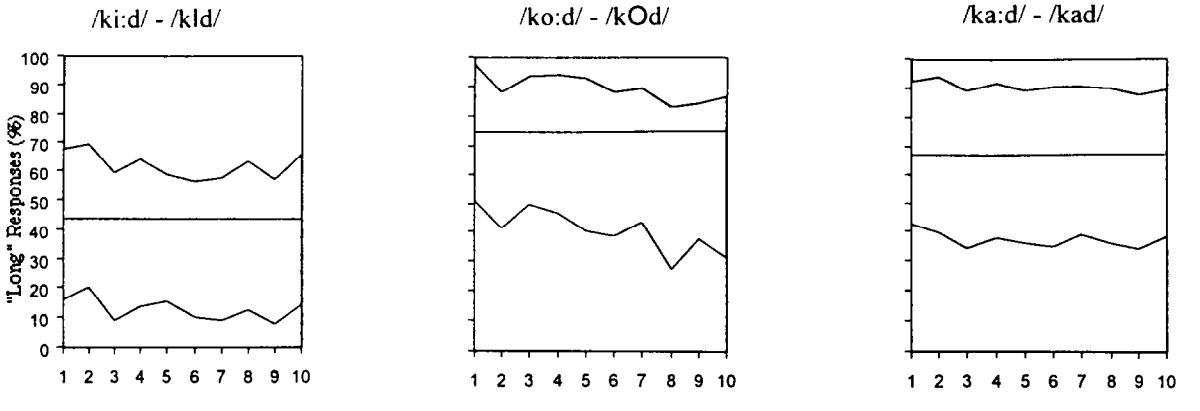

Figure 3: Confidence intervals (shaded areas) for percent long responses from the current study with adjusted postvocalic consonant duration are compared with the mean percent long responses (solid lines) for the corresponding vowel conditions in[4] for /ki:d/-/kId/, /ko:d/-/kod/, and /ka:d/-/kad/.