

Can V/C-ratio alone be sufficient for discrimination of V:C/VC: in Swedish? A perception test with manipulated durations

Bosse Thorén

Department of Linguistics, Stockholm University

Experience from second language learning and comparing regional varieties of Swedish, suggests that temporal organization, here defined as the ratio of vowel duration divided by consonant duration in stressed syllables, could be a stronger cue to V:C-VC: discrimination than the spectral difference between long and short vowel. 12 monosyllabic Swedish words were recorded and manipulated so, that the VC-sequence of each word had a duration of 600 ms, and given seven different V/C-ratios ranging from 0.1 to 1.4, based on values from the intact recordings. 20 native Swedish listeners judged for each stimulus whether it was a /V:C/-word or a /VC:/-word. The result shows that all vowels in the experiment were sensitive to temporal change, i.e. once the V/C-ratio was far enough from the original value, the listeners tended to interpret an original /V:C/-word as a /VC:/-word and vice versa.

1. Introduction

Out of the need to simplify linguistic description for pedagogic purposes, a basic prosodic model for Swedish has developed over the last two decades. This model often manifests itself in textbooks in Swedish as a second language in the following way: *Potatis är gott*.

In each prominent word, there is an underlining representing “the long sound in the stressed syllable”. This marking points out the prominent words in a sentence, the stressed syllables in the prominent words and finally the phonologically long sound in the stressed syllable, thus also promoting the quantity distinction by increasing or reducing the V/C-ratio.

The model has taken influence by Bruce’s (1977) model, where the quantity distinction – which requires a prolonged syllable – is the common denominator of all degrees of word prominence in Swedish. It also owes much to Bannert’s (e.g. 1986, 1995) many studies that have improved our understanding of what phonetic properties are crucial for speaking intelligible Swedish, and what properties are less important. His findings suggest that correct word stress is vital for making oneself understood in spoken Swedish. This view that stress and temporal organization is a basic property in spoken Swedish is corroborated by Fant & Kruckenberg (1994), who suggest that duration is the most reliable phonetic correlate to stress in Swedish.

Thorén (2001) tested digitally increased duration in phonologically long segments, as a means of improving Swedish with a foreign accent, and found that native Swedish listeners tended to judge increased duration as improved Swedish pronunciation, but some of the listeners had difficulties in discriminating manipulated and intact stimuli.

Elert (1964) has shown that at least central standard Swedish has a complementary vowel-consonant length relation in stressed syllables. This indicates that the V/C-ratio is a more reliable cue for /V:C/-/VC:/ distinction than absolute durations for long and short segments.

The distinction /V:C/ – /VC:/ is also accompanied by a spectral difference in the vowel, which is greater for some vowels and smaller for others. Duration and spectrum also interacts differently in different regional varieties of Swedish, and e.g. Reuter (1971) finds that Finland-Swedish has smaller spectral differences between long and short vowel than central standard Swedish, which is noticeable for the vowel phonemes /a/ and /ʌ/ in particular.

The main purpose of the present study is to test the power relation between duration and spectrum as cues to Swedish quantity, and to compare a method using complementary duration with studies using manipulation of vowel duration only, e.g. Hadding-Koch & Abramson (1964). Another aim is to scrutinize the suggestion of Hadding-Koch & Abramson (1964), that Swedish has "...a subset of eight vowel phonemes that combine with a phoneme of length to yield long counterparts and another subset of two vowel phonemes, /ʌ/ and /ø/, distinguished by quality."

2. Method

12 Swedish words, providing 6 minimal word pairs, differing only in quantity, were read by two native Swedish (Stockholm variety) speakers, one male and one female. The words were:

mat-matt [mɑ:t-mat:], vit-vitt [vi:t-vit:], hut-hutt [hʉ:t-höt:], våt-vått [vø:t-vøt:], tät-tätt [tɛ:t-tæt:] and söt-sött [sø:t-søt:].

Based on duration values from the readings of the two speakers, a duration of 600 ms was chosen as total duration of the VC-sequence for all stimuli in the experiment. The recording of the female speaker was resynthesized and manipulated in PRAAT (Boersma & Weenink 2001) and each of the 12 words were given V/C-ratios shown in table 1.

Table 1. Vowel durations, consonant durations and V/C-ratios for all test words.

V/C ratio in the Diagonal	Consonant duration						
	250	300	350	400	450	500	550
350	1,4						
300		1,0					
250			0,7				
200				0,5			
150					0,3		
100						0,2	
50							0,1

7 versions of each word made a total of 84 stimuli, that were presented in random order to 20 native Swedish listeners. Presentation was made with head phones from computer or CD-player. The listeners heard each stimuli once, and had to decide whether it was a /V:C/- or a /VC:/-word, by marking one of two words, e.g. *mat* or *matt*. There was no "uncertain" alternative. After the discrimination test, the listeners judged the naturalness of the pronunciation for each stimulus on a scale from 1 "very unnatural" to 10 "totally natural".

3. Result

The result of the discrimination test is captured in figure 1, and shows that all included vowel phonemes are influenced by the change in duration, and thereby V/C-ratio. Three out of six included vowel phonemes were subject to shift in /V:C/ - /VC:/ judgement by all the listeners. The word most resistant to temporal change was original “hutt”, which still made 11 listeners judge it as “hut” [hʉ:t] in its version with the greatest V/C-ratio (1.4).

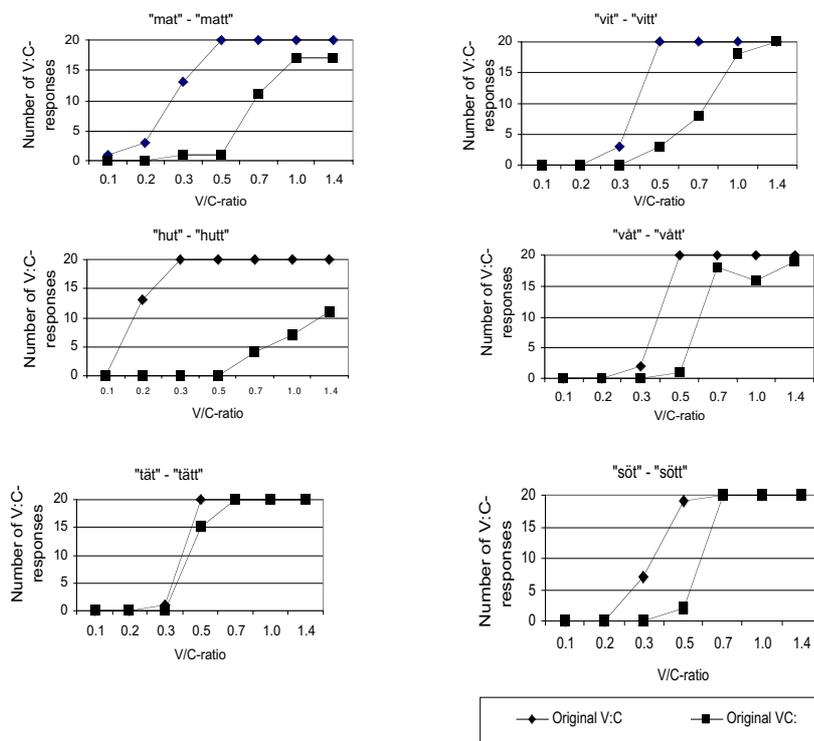


Figure 1. Number of listeners perceiving stimuli as /V:C/-word as a function of V/C-ratio. The leftmost graph always represents an original V:C-word and vice versa.

Two examples of naturalness judgements are presented in figure 2. One is for original “hutt” and the other for original “tat”, each of them representing an extreme in low and high sensitivity to temporal change without accompanying change in spectral properties. As could be expected, original “tat” [tɛ:t] is judged as highly natural both when perceived as /V:C/ [tɛ:t] and when perceived as /VC:/ [tɛ:t], with a slight lowering towards the middle and the edges. Original “hutt” [hɛ:t] is judged as natural only when having V/C-ratios 0.1 to 0.5 i.e. V/C-ratios similar to original version read by the Swedish speakers.

This is consistent with the traditional view, that /ɨ/ has greater spectral difference between long and short allophone than other Swedish vowel phonemes, and that /ɛ/ has minimal spectral difference between long and short allophone.

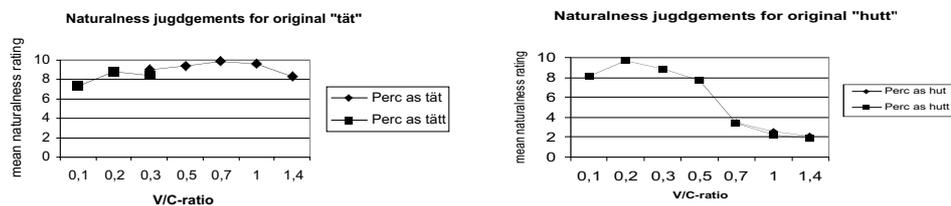


Figure 2. Naturalness ratings for original “tät” and “hutt”. The leftmost graph shows ratings for stimuli perceived as /VC:/ and vice versa.

4. Discussion

The study indicates that all of the included vowel phonemes use both duration and spectrum as cues for the quantity distinction, but in different proportions for various vowel phonemes. It does not support the suggestion of Hadding-Koch & Abramson (1964), that the /ʉ/ phoneme would hold an exceptional position in the Swedish vowel inventory, but it puts /ʉ/ at one end of a continuum of “spectrum dependence” and the /ɛ/ phoneme at the other end, requiring minimal change in spectrum for quantity distinction.

Considering the great changes in V/C-ratio that were needed especially for [ø] to be perceived as [ʉ], another conclusion is that temporal correlate to Swedish quantity is a necessary but not sufficient condition, which in turn has bearing on tutorial application. More similar experiments should be carried out with various Swedish regional varieties, to extract what is general Swedish quantity, and what is regional or local Swedish quantity.

5. References

- Bannert R. (1986) From prominent syllables to a skeleton of meaning: a model of prosodically guided speech recognition. *Working Papers, Department of Linguistics, Lund University* 29, 1-30.
- Bannert, R. (1995). Intelligibility in foreign accented Swedish: the effects of rhythmical and tonal features. *Reports from the Department of Phonetics, Umeå University. PHONUM*
- Boersma P. & Weenink D. (2001) Praat – a system for doing phonetics by computer. <http://www.fon.hum.uva.nl/praat/>
- Bruce, Gösta. (1977). Swedish word accents in sentence perspective. Gleerup, Lund.
- Elert, C-C. (1964). Phonologic Studies of Quantity in Swedish. Uppsala: Almqvist & Wiksell.
- Fant & Kruckenberg (1994). Notes on stress and word accent in Swedish, *KTH, Speech Transmission Laboratory, Quarterly Progress and Status Report 2-3*, 125-144.
- Hadding-Koch & Abramson. (1964). Duration versus spectrum in Swedish vowels: Some perceptual experiments. *I Studia Linguistica* 18. 94-107.
- Reuter M. (1971). Vokalerna i finlandssvenskan: En instrumentell analys och ett försök till systematisering enligt särdrag. *I Studier i nordisk filologi* 58, 240-249.
- Thorén B. (2001). Vem vinner på längden? Två experiment med manipulerad duration I betonade stavelser. *D-uppsats i fonetik. Institutionen för filosofi och lingvistik, avdelningen för fonetik*. Umeå universitet.