LOCUS-NUCLEUS RELATION AND VOT IN SPONTANEOUS AND ELICITED SPEECH

Diana Krull

Institute of Linguistics, Stockholm University, S-106 91 Stockholm, Sweden

ABSTRACT

CV-sequencies occurring in the spontaneous speech of five Swedes were compared to the same sequencies reproduced by each speaker in citation form words. Two acoustic characteristics are reported here: F_2 trajectories for C= /b, d, m, n, l/ and voice onset time (VOT) for C=/p, t, k/. Plotting F_2 at the CV boundary (locus) as a function of F_2 within the vowel (nucleus) resulted in steeper regression lines for spontaneous speech which was interpreted as a greater extent of contextual assimilation. Both locus and nucleus had also more central values in spontaneous speech. For VOT, no significant difference was found between spontaneous speech and citation form words, although both the duration of the consonantal closure and that the following vowel were considerably shorter in spontanaous speech.

1. INTRODUCTION

The present paper reports two studies comparing CV-sequences in spontaneous speech and in citation form words. The studies are a part of a larger investigation of phonetic variation in spontaneous Swedish[1].

The first study deals with the extent of contextual assimilation between the consonant and the vowel; the second addresses itself to comparisons of VOT.

1.1 Locus-nucleus relation of F2 and contextual assimilation

The size of the formant excursion from the CV boundary towards its "target" within the vowel has been shown to depend to a large extent on the duration of the vowel [7][4]. Moreover, the dimension of more or less clear pronounciation - "hypo" or "hyper" speech - is important: formant excursions have been shown to be larger in clear speech compared to neutral speech [10].

Plotting the locus frequency of a formant, e.g. F_2, as a function of its frequency within the vowel results in a linear function called the "locus equation", Eq.(1):

$$F_{2i} = k^{*}F_{2v} + c \quad (1)$$

where F_{2i} denotes the initial locus of the second formant, F_{2v} is the maximum or minimum within the vowel, and k and c are constants. Locus equations were first used by Lindblom [7] who demonstrated that the value of the constant k, i.e. the slope of the regression line, varies with consonant place of articulation. The slope also expresses the extent of contextual assimilation between the consonant and the vowel [5]. In the case of maximal assimilation, F_2 at the initial locus has the same frequency as in the middle of the vowel: k=1, and c=0. In the other extreme, the initial locus is invariant through all vowel contexts, k=0 and y has a constant value.

1.2 Experiment I

Five male speakers of Standard Central Swedish served as subjects. Natural continuous speech was obtained by asking each subject to relate a previously read short story, and to answering questions posed about the subject's work, travel, etc. The sessions were recorded and transcribed. Thereafter, word-initial CV-sequences with C=/b, d,m,n,l/ were located. F_2 was measured on wide band spectrograms at two points: "locus" at the CV boundary and "nucleus" at the maximum or minimum point within the vowel. If there was no minimum or maximum, the corresponding measurement was performed in the middle of the vowel. The words containing the CV-sequences used for measurement were then assembled in lists, separate for each speaker, who read the words separating them with pauses.

Plotting the locus as a function of nucleus resulted in slopes and y-intercepts given in Table I. It can be seen that, for a given place of articulation, the slope of the regression line is steeper for spontaneous speech, which can be interpreted as a sign of greater contextual assimilation. Of the possible factors influencing the extent of assimilation, the roles of lexical stress and phonological length were investigated, using only content words. The results showed that while there was relatively little change in slope with different phonological length, lexical stress caused a marked flattening of the slope. Higher k-values indicate that F_2 locus and nucleus were nearer each other in spontaneous speech. However, further investigation showed that the locus and and nucleus frequencies of the citation form words had not changed in a direction towards each other in spontaneous speech. Instead, both had moved towards a more central frequency value.

1.3 Discussion I

According to our interpretation of k-values in locus equations, there was always more contextual assimilation in spontaneous speech. Similar results have been obtained for French [1][2], Spanish and Catalan [11]. Both locus and nucleus frequencies were also more centralized in spontaneous speech. One probable reason for these differences is the usually shorter duration of the sequences in question and a resulting formant undershoot, i.e. the formant has not time to come near its target value [7][4]. Another reason for the difference may lie in the "hyper"-"hypo" dimension: the citation form words were usually more clearly pronounced than their spontaneous counterparts [10].

Table I. y-intercept and slope for the regression lines of initial locus vs. nucleus F2 in CV-sequences.

Speaker	OE	RL	JS	PT	ÅV
C=/d,n,l/					
	n=142	n=83	n=107	n=118	n=88
Reference					
y-intercept	1106	1193	1041	1033	870
slope	0.29	0.31	0.36	0.36	0.45
Spontaneous					
y-intercept	937	936	823	795	755
slope	0.32	0.39	0.44	0.47	0.51
C=/b,m/					
	n=64	n=52	n=88	n=64	n=36
Reference					
y-intercept	568	313	363	276	389
slope	0.58	0.74	0.70	0.79	0.72
Spontaneous					
y-intercept	407	244	261	160	179
slope	0.64	0.75	0.75	0.83	0.81

2. VOT IN SPONTANEOUS SPEECH IND IN CITATION FORM WORDS

Lisker and Abramson [8] compared VOT - the time between the stop release and the onset of the vocal cord vibrations - in isolated English words and in read sentences, showing that for a given CV-sequence VOT was considerably longer in isolated words. The role of several contextual features on VOT was studied, three of these were shown to have no effect: initial vs. non-initial position, utterance tempo, and vocalic environment. Stress, on the other hand, had a strong effect.

In Swedish, VOThas been shown to increase with stress in nonsense words [6][9]. In semantically meaningful sentences, VOT can be approximately doubled with the addition of emphatic stress [3]. The aim of this investigation is to study VOT in CV sequences in lexical words occurring in spontaneous speech, and in the same words words read in citation form.

2.1 Experiment II

The material consisted of two of the recordings described in section 1.1 above. This time, CV-sequences occurring in content words where located, C being a voiceless stop and V any vowel. For each CV-sequence, the duration of the stop gap, VOT, and the duration of the vowel were measured on wide band spectrograms. As in the previous experiment, word lists were prepared and read by the speakers.

The effect of four of these factors of possible influence on VOT are reported in this paper: (1) Stress (main and secondary); (2) place of articulation; (3) phonological length of the vowel and consonant; (4) the physical duration of the vowel and of the stop gap.

The results of the comparison revealed no significant difference between VOT in spontaneous speech and in citation form, although VOT tended to be slightly shorter in the isolated words (Table II). There was, however, a large difference in both in the duration of the stop gap and that of the vowel: both were much longer in citation form words.

Of the different factors whose influence on VOT was studied, only stress and place of articulation were shown to have a strong effect, both in spontaneous speech and in citation form words. Mean VOT was between 30% and 100% longer in stressed CVs than in corresponding unstressed syllables. In spontaneous speech as well as in citation form, the velar consonant had a longer mean VOT than the dentals and labials. The mean VOT for the dental consonant was in most cases longer than that of the labial. For both overlap in VOT between places of articulation as well as stressed and unstressed syllables.

According to t-tests, neither the phonological length of the vowel nor that of the consonant had a significant effect on VOT. There was, moreover, no significant correlation between the physical duration of the vowel and VOT. On the other hand, there was a weak but significant (p.05) negative correlation between the duration of the stop gap and VOT. (See [6] for detail).

Table II. The duration of the closure (stop gap), VOT, and V2 (in ms) in spontaneous speech and in citation form words. CV- sequences in word-initial position are not included.

	Closure	SD	VOT	SD	Vowel	SD	N
Speaker JS							
Unstressed CV							
spo	77	26	36	13	56	28	124
cit	163	47	34	12	135	41	
Stressed CV							
spo	63	17	46	13	79	37	28
cit	121	44	47	15	149	38	
Speaker PT							
Unstressed CV							
spont	90	34	35	12	56	30	101
cit	137	40	32	15	109	33	
Stressed CV							
spont	67	13	59	24	83	27	16
cit	72	15	56	9	129	34	

2.2 Discussion II

Lisker and Abramson's data [8] showed considerably longer VOT for words read in isolation than for words read in sentences. It was therefore surprising to find that in the present material VOT in isolated words tended to be slightly shorter than in spontaneous speech although the difference was not significant. The standard deviations for VOT were also approximately the same in both speaking styles, showing that the variation in VOT was not larger in spontaneous speech. The duration of the stop gap and that of the vowel, on the other hand, were both much longer in citation form words. There was also a greater variation in duration. The difference between the present results those of Lisker and Abramson [8] may be due either to language differences or to the fact that the connected speech in the American material was read text, while the Swedish material consisted of spontaneous speech. Possible differences in VOT between these two speaking styles have not been investigated.

FOOTNOTE

1 The project: Speech transforms - an acoustic database and phonetic and phonological rules for Swedish phonetics and phonology (Olle Engstrand, project director, Diana Krull, Björn Lindblom and Rolf Lindgren), supported by The Bank of Sweden Tercentenary Foundation, grant 86/109 and by The Swedish Board of Technical Development, grant 89-0027.

REFERENCES

[1] DUEZ, D. (1989). Second formant locus-nucleus patterns in spontaneous speech: some preliminary results on French. *PERILUS* X*,

[2] DUEZ, D. (1990). Some evidence on second formant locus-nucleus patterns in spontaneous speech in French. *PERILUS* XI*,.

[3] ENGSTRAND, O. (1983). Articulatory coordination in selected VCV utterances: A means - end view. (Ph.D. diss.). *Reports from Uppsala University Department of Linguistics* (*RUUL*) **10**, 76 ff.

[4] ENGSTRAND, O AND KRULL, D (1988). On the systematicity of phonetic variation in spontaneous speech. *PERILUS* VIII*, [5] Krull, D.

[5] KRULL, D. (1989). Second formant locus patterns and consonant-vowel coarticulation in spontaneous speech. *PERILUS* X*.

[6] KRULL, D. (1990). VOT in spontanous speech and in citation form words. *PERILUS* XI*,

[7] LINDBLOM, B. (1963). Spectrographic study of vowel reduction. *J. Acoust. Soc. Am.*, **35**, 1773-1781.

[8] LISKER, L. AND ABRAMSON, A. (1967). Some effects of context on voice onset time in English stops. *Language and Speech*, **10**, 1-28.

[9] LÖFQVIST, A. (1976). Closure duration and aspiration for Swedish stops. *Working Papers 13, Phonetics Laboratory, Lund University*, 1-39.

[10] MOON S.-J. AND LINDBLOM, B. (1989). Formant undershoot in clear and citation-form speech: A second progress report. *STL-QPSR* **1/1989**, Royal Institute of Technology, Stockholm.

[11] POCH OLIVÉ, D., FERNANDEZ-GUITIERREZ, N., AND MARTINEZ-DAUDEN, G. (1989). Some problems of coarticulation in CV stop syllables in Spanish and Catalan spontaneous speech. *Speech Research '89*, June 1-3, Budapest.

* Institute of Linguistics, Stockholm University