QUANTITY CONTRAST IN SWEEDISH KINEMATIC AND ACOUSTIC PATTERNS

Rudolph Sock and Anders Löfqvist

Institut de Phonétique de Strasbourg, 22 rue Descartes - 67084 Strasbourg, France and Haskins Laboratories, 270 Crown Street, New Haven-CT 06511 U.S.A.

ABSTRACT

The aim of this investigation is to find out if vowel quantity contrasts for Sweedish emerge in the acoustic VC and CV domains and also on the movement level, with increased speaking rate. Results obtained on both the acoustic and movement levels, suggest that quantity contrast is not only portrayed in the domains examined, but that it is also maintained across speech rate.

INTRODUCTION

Sweedish, is a language, that uses quantity for grammatical and lexical distinctions. It has a contrast between two lengths, and this contrast occurs only in lexically stressed syllables, where the vowel is either reffered to as being short or long. If the vowel is long, it is either word final or followed by a short consonant, VVC; if the vowel is short, it is followed by a long consonant, VCC [1]. It should, however, be noted that while quantity contrast in Sweedish may also be accompanied by additional correlates such as vowel quality and diphthongization, the importance of these factors is not easy to pinpoint [2].

The present investigation attempts to show how vowel quantity contrasts emerge, not only on the acoustic level but also, on the articulator movement level, specified by jaw, lower lip, tongue tip and tongue body vertical displacements. Moreover, the CV span will be examined to see how the complementary distribution of the quantities (*i.e.* short-long *vs.* long-short) contributes to the distinctivity of contrasts [2]. The effect of voicing on contrasts is also examined. Speech rate will be varied as a perturbing factor of the linguistic system; thus it would be possible to evaluate the robustness of the linguistic contrasts.

METHOD

One Sweedish subject produced CVCV words and non words that were embedded in a carrier sentence. Utterances were produced fourteen times at two speech rates, normal (conversational) and at a self-selected fast speaking rate, with all possible combinations of the vowels /i, a, u/ *vs.* /i:, a:, u:/ and the stop consonants /p, t, k/ *vs.* /b, d, g/. Thus there were 24 conditions in all: 2 speech rates (normal and fast) x 2 vowel lengths (short and long) x 3 consonant-types (bilabials, dentals and velars) x 2 consonant categories (voiced and unvoiced). Results reported here will only focus on the vowel /a/ context.

The movement data were recorded using the Haskins Laboratories three-coil transmitter system [3]; [4]. Receivers were attached to the upper lip, lower lip, jaw, tongue tip, tongue body, and tongue rear (locations of the tongue receivers are reffered to approximately). Additional receivers, placed on the bridge of the nose and on the upper incisors, were used for correction of head movements. Care was taken during each reveiver placement to insure that it was positioned at the midline with its long axis perpendicular to the sagittal plane[4].

To obtain instantaneous velocity, the first derivative of the position signals was calculated using a 3-point central difference algorithm. Velocity signals were smoothed using the same triangular window.

Based on articulatory events in the velocity signal, two cycles were identified in the movement of each articulator (jaw, lower lip, tongue tip and tongue body). These velocity cycles were determined, as the interval between successive negative or positive peaks associated, respectively, with the lowering or raising movement in the production of a vowel and a consonant. An oral opening phase, associated with the production of the vowel, was defined within the oral opening cycle, and an oral closing phase, associated with the production of the consonant, was defined within the oral closing cycle. Two acoustic cycles were also defined: one as the recurrence of the onset of a clear formant structure (*i.e.* in the VC domain), corresponds to the vocalic cycle; the other as the offset of a clear formant structure (*i.e.* in the CV domain), corresponds to the consonantal cycle. Acoustic phases were specified within the appropriate acoustic cycle, as the interval that presents a stable formant structure for the vocalic phase, and as the obstruent portion, for the consonantal phase. It is hypothesized that oral opening and vocalic phases would reveal quantity contrasts, while oral closing and consonantal phases would highlight concomitant consonantal differences.

RESULTS

Data processing is based on the percentage of time taken by each phase in its cycle. ANOVAs were performed on measured intervals as dependent variables and grouping factors *Quantity*, *Voicing*, speaking *Rate* and *Place* of articulation.

Quantity contrasts

Acoustic relative timing

In the Vocalic Cycle corresponding to the VC domain, and at a normal speaking rate, quantity differences for the unvoiced category emerge distinctly along the vocalic phase ($p<0.01$), while cycle or syllabic durations are comparable for VCCs *vs.* VVCs. These quantity differences are maintained with speech rate increase ($p<0.01$), due to the relative timing stability of the linguistic classes at around 58% of the cycle for long vowels and at around 35% for their shorter counterparts (see Figure 1, left panels). This finding is true also for the voiced category [2], although there is a general tendency for linguistic classes to converge as speaking rate increases; however, all classes remain distinct due to the combined effect of vocalic and syllabic differences ($p<0.01$).

Movement relative timing

In the articulator lowering peak velocity cycle — generally associated with the acoustic VC domain — and at a normal speaking rate, quantity differences for the unvoiced category also emerge distinctly along the oral opening phases ($p<0.01$); movement cycles still show comparable values for VCCs *vs.* VVCs. In their vertical lowering displacements, jaw, lower lip, tongue tip and tongue body opening movements for vowel production show clear-cut phasing patterns. Patterns obtained for jaw lowering, in the bilabial context, however, are more distinct than those observed for lower lip lowering, thus indicating that the jaw plays a more critical role in portrayiing vowel quantity contrasts. Figure 1(right panels) shows how short and long vowels emerge and differ as to the opening phases of the articulators; these differences are further maintained in fast speech largely by relative intraclass stability. These results are also valid for the voiced category.

Closure durations

Acoustic relative timing

In the consonant cycle corresponding to the CV domain, and at a normal speaking rate, closure differences for the unvoiced category emerge distinctly along the consonantal phase ($p<0.01$): the closure duration for the VCC category for /p, t, k/ take up around 65% of the CV domain, while that of the VVC category occupy, on an average, 45% of the cycle for the three consonant types. When speaking rate is increased, these consonant closure differences are maintained for /p, t/ (p<0.01) but tend to dissapear for /k/. The latter remark also applies to the voiced category, although closure differences become more critical in fast speech, especially for the lingual consonants.

Movement relative timing

In the articulator raising peak velocity cycle — generally associated with the acoustic CV domain — and at a normal speaking rate, closure differences for both the unvoiced and voiced categories do not emerge distinctly along the oral closing phases. The tendency, however, is for the VCCs to take up a higher closing phase percentage of the oral closing cycle.

Voicing contrasts

On the acoustic level, there is a systematic vowel phase difference, in the VC domain, between the voiced and unvoiced categories for both short and long vowels: the vowel, proportionally, is longer in the voiced context than in the unvoiced context at both speaking rates (p<0.01). In the CV domain, closure differences are less evident: unvoiced consonants tend, however, to take up a higher phase percentage of the closure cycle than their voiced counterparts.

On the movement level, the oral opening phase, associated with vowel production, does not show clear-cut patterns, even though the tendency is for vowels in the voiced context to have a longer opening phase than in the unvoiced context [4], especially for /p, t/. The oral closing phase, associated with consonant closure, does not show coherent pattern differences between the two categories.

In summary, these preliminary results show that, the effect of rate is significant as all cycles (syllables) are compressed with increase of speaking rate. However, linguistic contrasts are maintained by the relative stability of vowel (and associated oral opening) phases. The complementary consonantal differences also contribute to maintaining the linguistic contrast, although such differences are less resistant to speech rate, as revealed by the tendency for consonantal phases (and associated oral closing phases) to converge when the task becomes difficult. Systematic relative stability may be accounted for, to some extent, by biomechanical and aerodynamic constraints. However, acoustic and kinematic maintained differences that correspond to the different linguistic tasks are presumably constrained by the perceptual requirements of the linguistic code [5].

ACKNOWLEDGMENTS

This research was supported by NIH Grant DC-00865, from the National Institute of Deafness and Other Communication Disorders, the Fyssen Foundation, Esprit-BR Project 6975 — Speech Maps, and Grant P-55 from the Sweedish National Board for Industrial and Technical Development.

REFERENCES

[1] Elert C.C. (1964) Phonological studies of quantity in Sweedish (*Almqvist & Wiksell, Uppsala*).

[2] Engstrand O. Krull D. (1994) Durational correlates of quantity in Sweedish, Finnish and Estonian: cross-language evidence for a theory of adaptive dispersion. *Phonetica 51, 80-91.*

[3] Perkell J.S. Cohen M.H. Svirsky M.A. Matthies M.L. Garabieta I. Jackson M.T. (1992) Electromagnetic midsagittal articulometer systems for transducing speech articulatory movements. *J.A.S.A., 3078-3096.*

[4] Löfqvist A. Gracco V. (1994) Tongue body kinematics in velar stop production: influence of consonant voicing and vowel context. *Phonetica 51, 52-67.*

[5] Abry C. Orliaguet J-P. Sock R. (1990), Patterns of speech phasing. Their robustness in the production of a timed linguistic task: single *vs.* double (abutted) consonants in French. *European Bulletin of Cognitive Psychology 10,* 269-288.

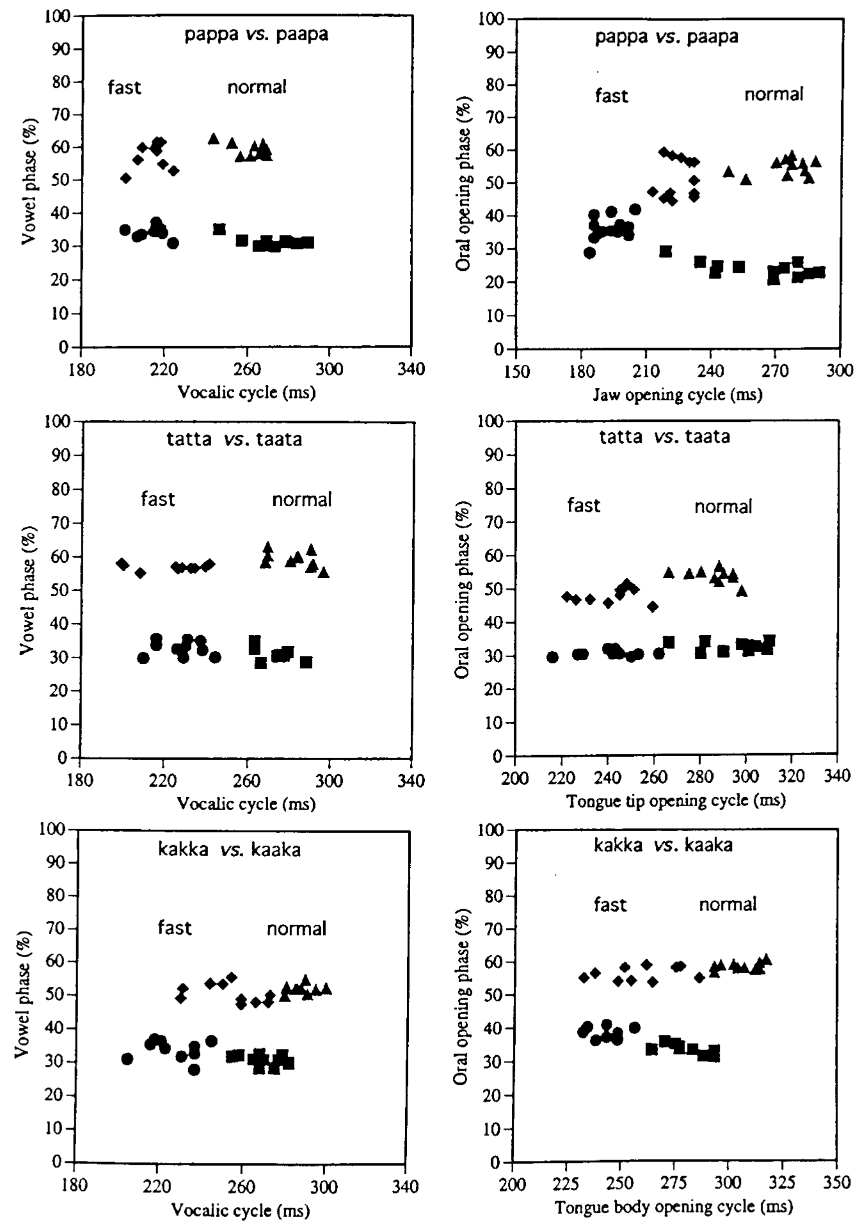

Figure 1. Scatterplots of vocalic patterns for VVCs in normal and fast speech (▲, ◆) and for VCCs in normal and fast speech (■, ●) on the acoustic level (left panels) and on the movement level (right panels). See text for details.