FLEXIBILITY AND STABILITY IN BILABIAL GESTURES: 2) EVIDENCE FROM CONTINUOUS SYLLABLE PRODUCTION

Pascal van Lieshout[*], Marjolein Hijl[‡], and Wouter Hulstijn[‡]
[*]*University of Toronto, Canada,*
[‡]*University of Nijmegen, The Netherlands*

ABSTRACT

In this paper the effect of syllable structure on bilabial coupling is investigated from a coordinative dynamics point of view. Data will be presented on four normal speaking subjects, who were asked to reiterate at a fast rate sequences of monosyllable and disyllable non-words, differing in syllabic structure. The dynamics of the coupling between upper and lower lip was measured in using continuous estimates of relative phase values. The results show a clear effect of syllable structure on the coupling of the lips while producing bilabial closing gestures. In particular, it was found that sequences in which cv-syllable structure was repeated, showed a marked difference in terms of lip coupling characteristics and associated speech rates compared to other sequences. The findings are discussed in terms of underlying dynamics of bilabial coupling, in which relative coordination and self-organization are key features.

1. INTRODUCTION

The syllable as a unit of speech production has received much attention in the (psycho) linguistic and motor control literature. It has been shown that repetition of the same syllable structure in non-word sequences, regardless of its phonetic context, facilitated verbalization [1]. In coordinative dynamics the syllable may act as a production unit where consonant movements are organized within the context of vowel-to-vowel cycles [2]. Furthermore, consonant and vowel productions are conceptualized as dynamically specified actions, in which individual articulatory movements are constrained to various, task-specific degrees (e.g. [3]). These constraints determine the amount of flexibility and stability that is available in the system [4].

In our companion paper [5], we looked at the variability and latency of bilabial relative timing for discrete syllables. The aim of this paper is to investigate the effect of syllable structure on the coupling dynamics between upper and lower lip in reiterated bilabial closure gestures. In particular, this study tests the hypothesis that the manner in which syllables are produced represents an emergent property of articulatory dynamics.

2. METHODS

2.1. Subjects

Four normal speaking male subjects (age range 23-25 years), all native speakers of Dutch, participated in the experiment. All subjects had normal hearing acuity, normal language and voice quality. They were volunteers paid for their participation.

2.2. Stimuli and procedures

In this experiment pronounceable but meaningless sequences of letter strings were used as stimuli. Each sequence belonged to one of seven possible categories (Table 1). For each category there were 4 different stimuli, repeated 2 times, giving a total of 56 sequences. A subject was instructed to repeat each of them as fast as possible for 10 seconds while maintaining intelligibility. Sequences contained one or more bilabial gestures ([p] sound). However, some sequences in categories 1 and 7 did not contain a bilabial gesture. These were not analyzed, since the main focus was on bilabial coordination.

As part of another experiment, not discussed here, subjects were required to produce each sequence in isolation in two separate blocks of trials prior to the experiment. This way potential problems in the pronunciation of the sequences could be identified and corrected.

Category	Vowel Context	Example
(1) CV.CV.CVC	Same	[pi.pi.təʁ]
(2) CVC.CVC.CVC	Same	[pip.pip.təʁ]
(3) CV.CV.CVC	Different	[pi.ta.kəʁ]
(4) CVC.CVC.CVC	Different	[pip.tap.kəʁ]
(5) CV.CVC.CVC	Same	[pi.pip.təʁ]
(6) CVC.CV.CVC	Same	[pip.pi.təʁ]
(7) CV.CVC.CV.CVC	Same	[pi.təʁ.pi.təʁ]

Table 1. List of categories that define syllabic structure

2.3. Instrumentation

Kinematic data associated with tongue, lips and jaw movements were collected using an electromagnetic midsagittal articulograph or EMMA system (AG100, Carstens Medizinelektronik GmbH). For this purpose, small sensor coils (about 2 mm in diameter) are attached to the surface of the tongue (on tongue blade, 1 cm behind apex, and on tongue body, 3 cm behind first tongue coil position), lips and lower jaw. For the lower jaw, a small thin mouthpiece was custom made for each subject (using Vinyl Polysiloxane; 3M Express STD) to fit the contours of the subject's lower incisors and gums. The jaw sensor was attached to this mouthpiece to ensure an accurate and reproducible position for transducing lower jaw movements.

Movement data were sampled at 400 Hz for each channel (both dimensions). Acoustic recordings, sampled at 16 kHz, were made simultaneously using the EMMA acoustic recording facilities. All data were processed and analyzed using HADES, a data processing and analysis software package [6].

2.4. Measurements

All movement data are normalized towards a common horizontal axis, as defined by EMMA measurements of the subject's occlusal plane. Coordination or coupling between upper and lower lip was measured using continuous estimates of relative phase [7]. Average relative phase values and their standard deviations were calculated using circular statistics [8]. An illustration of this measure is shown in Figure 1.

From the acoustic recordings, the middle four seconds of a ten-second trial were selected. The number of syllables within this interval as well as the interval-durations between the vowel peaks of the individual syllables were calculated. The latter measure, denoted as mean syllable duration, was used for further analysis. Contrasts of interest were tested using Watson's F-test for two circular means. Alpha-levels for significance were set at 0.05, adjusted for the number of contrasts using a Bonferoni correction.

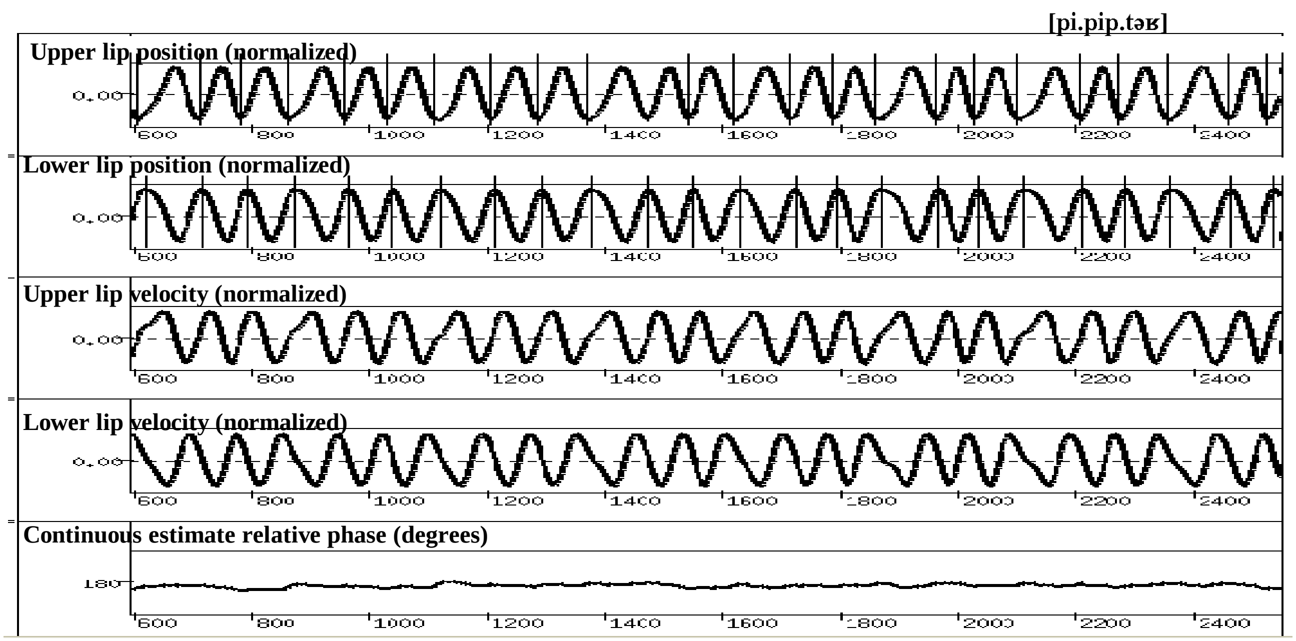

Figure 1. Illustration of continuous estimate of relative phase for normalized upper and lower lip data

3. RESULTS

A first analysis (per subject) was done on differences in relative phase as a function of vowel context. According to Sevald et al. [1], such a difference should not have a significant effect on the ease of articulation for a given syllable structure. Indeed, the results showed no significant difference for any of the subjects between sequences of categories 1 and 3 and sequences of categories 2 and 4. Furthermore, no differences were found for stimuli belonging to the categories 1 and 7. Therefore, the data from categories 1, 3 and 7 were combined as well as the data from categories 2 and 4. Thus, the principal contrasts of interest are between stimuli with a congruent cv (1+3+7) and cvc (2+4) pattern and discongruent, e.g. cv-cvc, (5+6) patterns.

Table 2 lists the average bilabial relative phase values (Phi) and average standard deviations for relative phase within trials (SD Phi) for each subject. The mean syllable durations (Duration) found for the cv-congruent (cv), cvc congruent (cvc), and incongruent (Inc) conditions are added in the third column. For the average relative phase values, all subjects showed a significant difference between the cv and cvc conditions, as well as between the cv and Inc conditions. However, differences between the cvc and Inc conditions were only found to be significant for s03. In general, the cv condition showed a more symmetric pattern (closer to 180 degrees) then the other conditions. The differences for average syllable durations paralleled the effects for relative phase in that mean syllable durations for the cv-condition were much shorter then for the other conditions. Regarding the standard deviations for relative phase (based on individual trial data), the effects were somewhat mixed. Subjects 1 and 2 showed significant differences between cv and cvc, as well as between cvc and Inc conditions. However, such effects were not found for subjects 3 and 4.

In Figure 2 the individual data for relative phase are plotted against the individual data for mean syllable duration. A few observations can be made. First, cv and cvc conditions are clearly distinguished in terms of their combination of relative phase and mean syllable duration values. Part of the latter effect is obviously related to the difference in number of segments [9]. Furthermore, the Inc condition seems to be characterized by values that can take either side, that is, they may fall within the region of the cv-condition, or they fall within the region of the cvc condition. One could say that in producing bilabial gestures

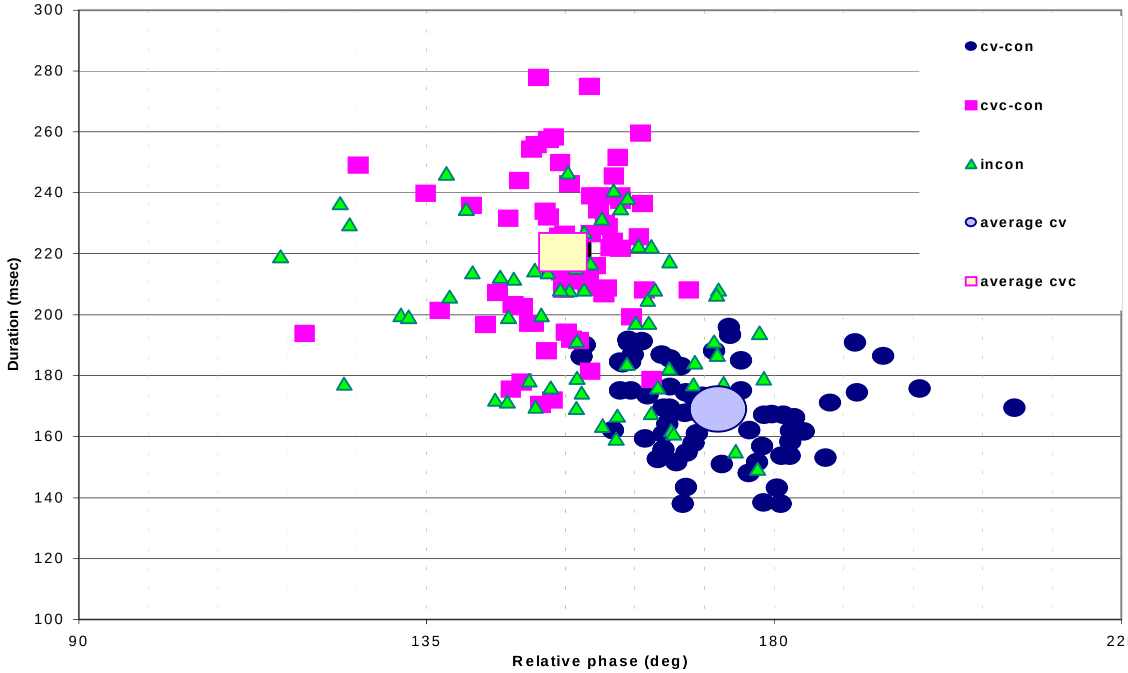

Figure 2. Relationship between relative phase and mean syllable durations for all subjects for cv-congruent (cv-con), cvc-congruent (cvc-con) and all incongruent (Incon) conditions

in these sequences, this articulatory system is attracted by one of two possible states or attractors. For a given setting, the (linear) relationship between speech rate and relative phase value is weak, both for the cv (R^2 = 0.05) and cvc (R^2 = 0.01) condition.

	Phi	SD Phi	Duration
S01			
Cv	176.68 (14.7)	44.56 (6.5)	184.38 (8.1)
Cvc	152.19 (8.6)	22.96 (6.7)	237.40 (23.3)
Inc	147.31 (16.3)	37.55 (15.9)	221.04 (16.2)
S02			
Cv	171.44 (10.8)	42.04 (1.9)	171.8 (10.2)
Cvc	151.06 (10.9)	26.74 (2.7)	234.2 (17.3)
Inc	153.03 (11.2)	42.13 (3.0)	207.6 (21.1)
S03			
Cv	174.59 (7.0)	20.43 (7.9)	149.8 (7.5)
Cvc	155.41 (6.0)	24.43 (6.5)	205.3 (20.3)
Inc	165.19 (9.2)	32.17 (11.4)	171.9 (12.8)
S04			
Cv	168.21 (6.3)	21.13 (8.7)	169.3 (10.7)
Cvc	152.09 (6.5)	19.27 (8.9)	205.1 (21.7)
Inc	156.30 (12.3)	25.46 (10.2)	193.2 (19.4)

Table 2. Mean relative phase values (Phi), mean standard deviations of relative phase across individual trials (SD Phi), and mean syllable durations (Duration) for cv-congruent (cv), cvc-congruent (cvc), and incongruent (Inc) conditions (between () standard deviations across repetitions)

There is a complicating factor in this study. In repeating sequences at a high speech rate with two bilabials following each other very closely in time (as in [pip.pi.təʁ]), it is possible that the two bilabials are reduced to a single gesture, thereby creating a different type of sequence ([pi.pi.təʁ]).

A closer look at some of the movement patterns for this type of sequences indeed showed examples of cluster reduction. For example, in Figure 3 it is shown that this subject produces mostly a simplified sequence ([pa.pa.təʁ]) instead of the intended more complex sequence [pap.pa.təʁ], although the subject does produce it sometimes as seen in the right part of Figure 3. Also note, that at the moment where a [t] sound is produced, the lower lip moves upward.

Even though all lower lip movements are (linearly) corrected for jaw movements, there seems to be some influence of the upward jaw movement present. The most interesting part, however, is the fact that the upper lip makes a corresponding downward movement (but to a lesser extent as for a target lip closure).

This upper lip movement does not appear to be a simple mechanical or reflex based reaction to the lower lip movement, since the upper lip movement starts at the same time or in some cases even precedes the onset of the lower lip movement (see ellipses in Figure 3). That is, the two lips behave like a unit, maintaining their coupling, even if there is no acoustic or phonological target for such coordinated lip movements as in [t] production.

Additional analysis, the details of which are left out due to space limitations, revealed that in general there is a clear separation in terms of relative phase and speech rate characteristics for cv-congruent constructs and other syllable constructs. This finding seems in line with the claim that the CV syllable has a special status in speech production ([10]). Furthermore, it was found that for the non-cv congruent structures, the number of bilabial gestures within a particular sequence influenced the variability of the relative phasing. More bilabials in a sequence reduced the variability of the bilabial coupling.

4. DISCUSSION AND PRELIMINARY CONCLUSIONS

Looking at the coordination between upper and lower lip in verbal sequences that differ in syllabic structure, it was found that:

a. if the second syllable of the sequence repeats a cv - structure, the coupling between the lips
   - tends towards symmetry (180^0)
   - is relatively variable ($SD > 25^0$)
   - is associated with relatively faster speech rates (around 170 ms/syllable)

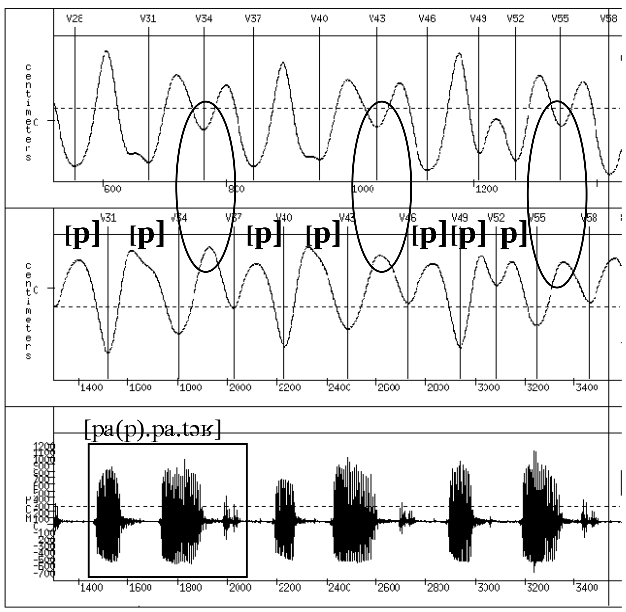

Figure 3. Example of sequence with mixed occurrences of double (intended) and simplified bilabial clusters

b. if the sequence represents the other conditions (cvc-congruent or any discongruent condition), the coupling between the lips
   - tends to be more asymmetric (around 150^0)
   - is relatively stable ($SD < 25^0$)
   - is associated with relatively slower speech rates (around 200 ms)
   - shows a decrease in relative phase variability with an increase in the number of bilabial gestures in the sequence

These characteristics were found to be rather consistent across different speakers and suggest that they act as indices of relatively stable solutions ("attractors") for bilabial coupling. However, it should be emphasized that interarticulatory coupling is not characterized by invariant (point-attractor) targets. Rather, the coordination between the lips shows a more flexible type of organization, as we would expect from a truly dynamical system [4,11].

We also noticed that in a number of trials subjects added bilabial movement sequences, the main purpose of which may be to maintain the integrity of the phase coupling between the lips. Since the added gesture does not have a valid acoustic or phonological target, it may indicate that speech movements are coordinated as coupled (non-linear) oscillators beyond the time-window of a single discrete gesture. Furthermore, the interaction between the coupled articulators determines the type and stability of the coupling, which is a trait of self-organization. For a given structure speech rate and relative phase seems to be rather independent from each other. Further analysis on the time-series of relative phase data are in progress in order to gain more insight in the underlying dimensionality of these signals, and the way the dynamics of the speech motor system may differ between normal speakers and people who stutter.

ACKNOWLEDGMENTS

The Dutch Organization for Scientific Research (NWO) is gratefully acknowledged for funding this project.

REFERENCES

[1] C.A. Sevald, G.S. Dell, and J.S. Cole, Syllable structure in speech production: Are syllables chunks or schemas? *Journal of Memory and Language*, vol. 34, pp. 807-820, 1995.

[2] C.P. Browman and L. Goldstein, Some notes on syllable structure in articulatory phonology *Phonetica.*, vol. 45, pp. 140-155, 1988.

[3] S. Shaiman and R.J.J. Porter, Different phase-stable relationships of the upper lip and jaw for production of vowels and diphthongs *J.Acoust.Soc.Am.*, vol. 90, pp. 3000-3007, Dec, 1991.

[4] J.A.S. Kelso. *Dynamic patterns. The self-organization of brain and behavior*, Cambridge, MA: A Bradford Book (MIT Press), 1995.

[5] P.J. Alfonso and P. van Lieshout. Flexibility and stability in bilabial gestures: 1) Evidence from discrete syllable production. In: *Proceedings of the Fourteenth International Congress of Phonetic Sciences*, Anonymous1999.

[6] P.E. Rubin. HADES: A case study of the development of a signal analysis system. In: *Behavioral aspects of speech technology: Theory and applications*, eds. R.Bennett, S.L.Greenspan, and A.Syrdal. Elsevier Science Publisher, 1995.

[7] J.A.S. Kelso, E. Saltzman, and B. Tuller, The dynamical theory in speech production: Data and theory *Journal of Phonetics*, vol. 14, pp. 29-60, 1986.

[8] E. Batschelet. *Circular statistics in biology*, London: Academic Press, 1981.

[9] W.R. Tiffany, The effects of syllable structure on diadochokinetic and reading rates *J.Speech Hear.Res.*, vol. 23, pp. 894-908, Dec, 1980.

[10] B. Tuller and J.A. Kelso, The production and perception of syllable structure *J.Speech Hear.Res.*, vol. 34, pp. 501-508, Jun, 1991.

[11] P. van Lieshout, W. Hulstijn, P.J. Alfonso, and H.F. Peters. Higher and lower order influences on the stability of the dynamic coupling between articulators. In: *Speech production: Motor control, brain research and fluency disorders*, eds. W. Hulstijn, H.F. Peters, and P. van Lieshout. Amsterdam: Elsevier Science Publishers, 1997.pp. 161-170.