



HAL
open science

An articulatory basis for the Labial-to-Coronal effect: /pata/ seems a more stable articulatory pattern than /tapa/

Amélie Rochet-Capellan, Jean-Luc Schwartz

► **To cite this version:**

Amélie Rochet-Capellan, Jean-Luc Schwartz. An articulatory basis for the Labial-to-Coronal effect: /pata/ seems a more stable articulatory pattern than /tapa/. *Journal of the Acoustical Society of America*, 2007, 121 (6), pp.3740-3754. <10.1121/1.2734497>. <hal-00157938>

HAL Id: hal-00157938

<https://hal.science/hal-00157938v1>

Submitted on 27 Jun 2007

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

1 An articulatory basis for the Labial-to-Coronal effect:

2 /pata/ seems a more stable articulatory pattern than /tapa/

3
4
5 Amélie Rochet-Capellan and Jean-Luc Schwartz

6
7
8
9 *Institut de la Communication Parlée, UMR5009 CNRS/INPG/Université Stendhal, 46*
10 *Avenue Félix Viallet, 38031 Grenoble Cedex 01, France*

11
12
13 Manuscript #06-03269

14 First online submission: 26/05/2006

15 Second online submission: 1/02/2007

16
17
18
19
20
21 Abbreviated title: An articulatory basis for the Labial-to-Coronal effect.

22

23 This paper investigates the coordination between the jaw, the tongue tip and the lower
24 lip during repetition with rate increase of Labial-to-Coronal (L_aC_o) Consonant-
25 Vowel-Consonant-Vowel disyllables (e.g./pata/) and Coronal-to-Labial (C_oL_a) ones
26 (e.g. /tapa/) by French speakers. For the two types of disyllables: (1) the speeding
27 process induces a shift from two jaw cycles per disyllable to a single cycle; (2) this
28 shift modifies the coordination between the jaw and the constrictors and (3) comes
29 with a progression towards either a L_aC_o attractor (e.g. (/pata/ or /tapa/) \rightarrow /patá/ \rightarrow
30 /ptá/) or a C_oL_a one (e.g. (/pata/ or /tapa/ \rightarrow /tapá/ \rightarrow /tpá/). Yet, (4) the L_aC_o attractor
31 is clearly favored regardless of the initial sequencing. These results are interpreted as
32 evidence that a L_aC_o CVCV disyllable could be a more stable coordinative pattern for
33 the lip-tongue-jaw motor system than a C_oL_a one. They are discussed in relation with
34 the so-called LC effect that is the preference for L_aC_o associations rather than C_oL_a
35 ones in CV.CV disyllables in both world languages and infants' first words.

36
37
38
39
40
41
42 PACS numbers: 43.70Aj, 43.70Fq

43 1. Introduction

44 The analysis of CV.CV disyllables in world languages lexicons reveals that Labial-to-
45 Coronal (L_aC_o) sequences (e.g. /pata/) are about 2.5 times more used than Coronal-to-Labial
46 (C_oL_a) ones (e.g. /tapa/) (Locke, 1983; MacNeilage *et al.*, 2000a, 2000b; Rousset, 2003). This
47 asymmetry, also displayed in infants' first words in the course of ontogeny (Ingram, 1974; Locke,
48 2000, MacNeilage *et al.*, 2000a, 2000b), is known as the "LC effect". The existence of the same
49 trend in infant and adult speech lead MacNeilage and Davis (2000b) to propose a "substance-
50 based" explanation of the LC effect in the framework of the "Frame-Content Theory" of speech
51 development (MacNeilage, 1998). Yet, there are some problems with their explanation, letting
52 space for further hypotheses. The present investigation is an attempt to test an original
53 explanation of the LC effect, based on motor principles likely to intervene both in speech
54 development and in on-line adult speech production.

55 1.1. The "simple first" hypothesis

56 MacNeilage and Davis (2000a, 2000b) proposed that the LC effect might result from the
57 conjunction of two principles in the course of speech development: (1) labial proto-consonants
58 (or closants in the Frame-Content terminology) would be easier to produce by infants than
59 coronal ones, and (2) in order to deal with the complexity of the mental lexicon access required in
60 the production of their first words, infants would focus on easier forms. Thus, they would select
61 sequences of actions beginning by the simpler gesture, supposed to be the labial one. The
62 preference for L_aC_o disyllables in development would then have been preserved in adults'
63 languages.

64 The assumed greater simplicity of labial gestures is one component of the Frame-Content
65 theory of speech motor control development (MacNeilage, 1998; MacNeilage and Davis, 2000a,
66 2000b). This theory claims that speech production begins in babbling with phonation associated
67 with repeated vertical oscillations of the jaw. Apart from vocal tract pre-settings, the other
68 articulators would stay stable all along the jaw cycles. Thus, besides their alimentary function,
69 jaw cycles would induce alternation of closing and opening patterns (*closants* and *vocants*) and
70 constitute the *frame* of speech. Then, the speech segmental *content* (independent control of
71 consonants and vowels inside the frame) would progressively emerge from the development of
72 the central and peripheral motor control of the other articulators. According to MacNeilage and
73 Davis, for neutral vocal tract pre-setting, that is purely passive tongue and lips, jaw upwards
74 gestures would induce a labial closure of the vocal tract and so, the production of labial closants
75 (“pure frames”). On the contrary, coronal closants would require an active motion of the tongue
76 in pre-setting (“fronted frames”). Therefore, labial closures would be easier to realize than
77 coronal ones. However, the “simple first” explanation suffers in our view from several
78 weaknesses. Firstly, it is a bit ad hoc. Secondly, the “simpler” nature of labial closants in
79 babbling has been questioned by Vilain *et al.* (1999) in a study exploiting articulatory models of
80 the vocal tract. They compared three models with different morphologies – coming from the
81 statistical analysis of cineradiographic data recorded from three different adult speakers. They
82 simulated pure frames on the three models, by applying jaw upward commands, all other
83 articulators being kept passive, and they obtained different kinds of closures. In one model the
84 jaw raise resulted in closing the lips first, producing a labial contact, while in another one it first
85 resulted in a closure inside the vocal tract, producing a coronal contact. In the last model, a labial

86 and a coronal contact were obtained almost in synchrony. Therefore, both /ba/, /da/ and /bda/
87 could be pure frames according to the vocal tract morphology. This could explain why analyses
88 of babbling inventories do not display a preference for labial patterns over coronal ones (Locke,
89 1983).

90 Another problem concerns the extrapolation from infants to adults. Indeed, there is no
91 reason to believe that labial plosives would stay simpler than coronal ones in adult speech, both
92 involving a specific articulatory control superimposed on the jaw, respectively for the lips in
93 labials and for the tongue in coronals (Munhall and Jones, 1998). Finally, the “simple first”
94 hypothesis tends to consider the articulatory control of L_aC_o and C_oL_a utterances as a sequence of
95 independent phonetic segments while it might be better characterized as coordinative structures
96 (Kelso *et al.*, 1986; Nittrouer, 1991).

97 1.2. An alternative hypothesis based on jaw-tongue-lips coordination

98 It is widely assumed that some coordinative patterns are more stable and economic than
99 others and that the repetition of a multi-effectors gesture with rate increase would attract the
100 system toward its most stable coordinative mode. This has been first established for inter-limb
101 coordination (Hoyt and Taylor, 1981; Haken *et al.*, 1985) and then, extended to articulators’
102 coordination in speech production (Kelso *et al.*, 1986; Lindblom, 1990a; Kelso, 1995). Such
103 criteria of stability and economy have been introduced in substance-based criteria shaping the
104 sound systems of human languages (e.g. Stevens, 1989; Lindblom, 1990b). In this vein, it has
105 been shown that rate increase induced repeated VC syllables to switch towards cycled CV ones
106 (Stetson, 1951; Tuller and Kelso, 1991; de Jong, 2001). This led the authors to conclude that CV

107 sequences are more stable than VC ones, which could explain the preference for CV syllables
108 over VC ones in language inventories (Redford and Diehl, 1999; Rousset, 2003).

109 These studies provide an interesting framework for attempting to anchor the LC effect in
110 the properties of the articulatory motor system rather than in a pure developmental framework.
111 More precisely, we assume that a labial constriction followed by a coronal one could be a better
112 coordinative pattern for the lip-tongue-jaw system than the coronal-then-labial one. This
113 assumption is inspired from Sato *et al.* (2006). The authors supposed that in CCV or CVC
114 sequences involving a labial and a coronal consonant, there was a trend to anticipate the coronal
115 constriction during the labial one, rather than the inverse. Hence, lips and tongue were more in-
116 phase in /psə/ than in /spə/. The reason is that in /psə/, the tongue tip may be prepared for /s/
117 during the labial closure, so that the /p/ release is almost synchronous with the tongue-driven
118 onset of the next /s/. On the contrary, in /spə/, the lips must stay open during the production of the
119 /s/, before labial closure followed by the /p/ release, so that the tongue-lip synchrony is difficult
120 to achieve (see also Byrd, 1996; Zsiga, 1996; Surprenant and Goldstein, 1998, for compatible
121 data in English; and Chitoran *et al.*, 2002 for Georgian). This articulatory asymmetry between the
122 two sequences of consonantal gestures could be a part of the LC effect explanation in both
123 language development and world languages lexicons.

124 1.3. An experimental scenario

125 The presumed greater stability of the coordination in L_aC_o compared with C_oL_a could be
126 investigated through the speeding paradigm. The slow regular repetition of C_1VC_2V disyllables
127 should be characterized by a regular alternation of C_1 and C_2 gestures over time. On the contrary,

128 rate increase could lead to asymmetries, the C_2 constriction being either closer to the previous or
129 to the next C_1 one. This would induce the reduction of either the vowel after C_1 or after C_2 . The
130 French language, under study here, displays a stress on the word last syllable. Therefore, in the
131 repetition with rate increase of a CVCV disyllable, French speakers should rather reduce the first
132 vowel, and a CCV re-syllabification could be expected. Hence, two hypotheses may be
133 contrasted. In a first one, there would be no asymmetry between L_aC_o and C_oL_a coordination, and
134 the French speakers would preserve the original disyllable prosody through speeding. In this first
135 hypothesis, L_aC_o CVCV sequences would evolve toward a L_aC_o attractor (e.g. /pata/ → /patá/ →
136 /ptá/) while C_oL_a CVCV sequences would evolve toward a C_oL_a CCV attractor (e.g. /tapa/ →
137 /tapá/ → /tpá/). Yet, if, as we expect, the L_aC_o coordination is more stable than the C_oL_a one, this
138 would attract the system towards a L_aC_o attractor regardless of the initial CVCV sequence (e.g.
139 (/pata/ or /tapa/) → /patá/ → /ptá/). Hence, in this second hypothesis, the repetition with rate
140 increase of both L_aC_o and C_oL_a CVCV disyllables would evolve towards a L_aC_o attractor. The
141 first experiment, based on acoustic measurements, aimed at selecting which of these two
142 assumptions is true.

143 In a second experiment based on articulatory measurements, we studied how the
144 articulatory coordination between jaw, tongue and lips would be reorganized in order to subtend
145 the progression from one constriction per jaw cycle (e.g. /pata/) towards two constrictions per jaw
146 cycle (e.g. /ptá/). A prediction was that the two constrictions would be realized on a single jaw
147 cycle before the fading of the vowel was completed (e.g. /patá/). In other words, the variation of
148 the articulation place between the first and the second constriction would allow the production of
149 two syllables on a single jaw cycle, while most studies of jaw movements associate the jaw cycle

150 with the CV syllable (Nelson *et al.*, 1984; Ostry and Flanagan, 1989; Perkell and Zandipour,
151 2002; Perkell *et al.*, 2002).

152 2. An acoustic study of LC and CL stability

153 The aim of this first study is to test if L_aC_o sequences are more stable than C_oL_a ones in a
154 speeding paradigm, using an acoustic measure of vowel reduction.

155 2.1. Method

156 2.1.1. *Speakers and phonetic material*

157 28 native French speakers with no speech or hearing problem and naive as to the purpose
158 of the study participated in the experiment. The phonetic material consisted in six CVCV
159 disyllables: three L_aC_o (/pata/, /pasa/ and /fata/) and their three C_oL_a counterparts (/tapa/, /sapa/
160 and /tafa/). This choice aimed at contrasting L_aC_o and C_oL_a sequences while controlling for a
161 possible “plosive-to-fricative” effect. Indeed, in the framework of “sonority” principles (Selkirk,
162 1984; Clements, 1990) it could be supposed that utterances containing a fricative and a plosive
163 would switch towards plosive-fricative attractors (e.g. /tfa/, /psa/), rather than the L_aC_o attractors
164 /pta/, /psa/ and /fta/ expected in the framework of our basic assumptions. The /a/ vowel was
165 selected to provide an open configuration, mainly involving the jaw, with basically no strong
166 requirements on the lips and the tongue. The six disyllables were randomly arranged inside three
167 lists. The three lists were presented to each speaker in different orders. Hence each disyllable was
168 presented three times to each speaker. The test began with three items provided as training
169 stimuli (/vada/, /daba/ and /dava/).

170 2.1.2. *Procedure*

171 In order to simplify post-treatments the experiment was assisted by a computer program
172 that allowed to record utterances and to display a kind of visual metronome. The participants
173 were sitting in front of a computer connected to a microphone. The instruction was to repeat the
174 disyllable displayed on the screen at an increased and then decreased rate following as much as
175 possible the rhythm of a visual flash. The aim of this flash was not to precisely control the
176 speakers' rate but rather to give global timing marks in order to homogenize acceleration and
177 deceleration phases between speakers and to attempt to drive the speakers to their limit rate. It
178 was mentioned that the flash would reach rates impossible to follow and the instruction was to
179 maintain a rate as high as possible during this period until the decelerating phase. Moreover,
180 speakers were encouraged to go on without interruption even if they perceived any kind of
181 transformation of the uttered sequence.

182 The experiment began by the three practice trials. Then, the three lists were successively
183 displayed. Each trial started with the presentation of the CVCV sequence to repeat, displayed in
184 the middle of the screen. Speakers initiated the flash pressing the “space” key. A blue square took
185 the place of the disyllable on the screen for one second, followed by the flash. The flash was an
186 alternation of a black and a white square. The instruction was to produce one syllable on the
187 black flash and the other on the white one. Duration of each square presentation started at 300 ms
188 and progressively decreased to reach 125 ms at the middle of the acceleration phase and 50 ms at
189 the end of the acceleration phase (8 s). The progression was shaped in order to obtain a large
190 amount of rapid productions. The timing of the deceleration phase was the symmetric of the
191 acceleration one. Thus, each trial record lasted 16 s. Speakers sometimes took a breath inside the

192 whole acceleration-deceleration process, but this occurred seldom and at random positions inside
193 the 16 s utterance. Then the speaker spontaneously resumed the process at about the same tempo
194 he/she had reached before the breath.

195 *2.1.3. Hypotheses and measurements*

196 The working hypothesis is that both L_aC_o and C_oL_a repetitions would evolve towards a
197 L_aC_o CVCV sequencing until eventually reaching a L_aC_o CCV consonantal cluster. For example,
198 if /pata/ is indeed more stable than /tapa/, the speeding process should be: (/pata/ or /tapa/) →
199 /patá/ → /ptá/. The contrasted prediction is that, if there were no asymmetry between L_aC_o and
200 C_oL_a utterances in terms of coordinative structure, the French prosodic pattern with a stress on the
201 final syllable of a word should drive the behavior with e.g. /pata/ → /patá/ → /ptá/ and /tapa/ →
202 /tapá/ → /tpá/. Finally, both trends could be observed, with a L_aC_o vs. C_oL_a asymmetry,
203 superimposed on an effect of the initial uttered sequence. These predictions were tested by
204 prosodic measurements based on vowel intensity. Acoustic energy was continuously estimated,
205 within the Praat software¹ using a 42.6 ms Kaiser-20 window (sidelobes below -190 dB).
206 Maxima and minima of energy curves were automatically detected and then hand-selected as
207 vowels and consonants (Figure 1). When a speaker's production evolved towards a single
208 consonantal cluster (e.g. /ptá/), the labeling rule was to mark the “vanished” vowel and the two
209 consonants around as a same event (e.g. Figure 1, right). Speakers' errors such as repetition of a
210 single syllable or pauses were removed. From these data, two parameters were computed for each
211 CVCV or CCV sequence correctly produced (now referred as “utterances”): (1) the utterance
212 duration and (2) the variation of intensity between the two vowels, that is the difference between

213 the intensity of the vowel after the coronal consonant (V_{C_o}) and the intensity of the vowel after
214 the labial consonant (V_{L_a}):

$$215 \quad \Delta I = I(V_{C_o}) - I(V_{L_a})$$

216 The operational hypothesis was that ΔI would be positive for both the $L_a C_o$ and the $C_o L_a$
217 utterances for fast productions, with very high positive values indicating a shift towards a single
218 $L_a C_o$ CCV syllable (e.g. from /pata/ or /tapa/ to /ptá/, as in Figure 1).

219

220 ----- Figure 1 -----

221

222 2.2. Results

223 The analyses excluded two speakers, one because he did not manage to alternate the labial
224 and the coronal constrictions at a fast rate and the other for recording problems. Apart from local
225 errors, all the 26 remaining speakers' utterances respected a "... $L_a V C_o V L_a V C_o V$..." sequencing,
226 L_a and C_o respectively being the labial and the coronal consonants of the target CVCV sequence,
227 and "V" the vowel /a/ or nothing.

228 2.2.1. Evolution of ΔI across rate: global tendencies

229 The plot of ΔI values against utterance durations for all the speakers' $L_a C_o$ and $C_o L_a$ utterances
230 shows very similar patterns for the three $L_a C_o$ and $C_o L_a$ pairs. Hence, only the graphs for /pasa/
231 and /sapa/ are displayed on Figure 2 (top), as a representative example. Globally and for both the
232 $L_a C_o$ (left) and $C_o L_a$ (right) groups, utterance durations range from 100 to 700 ms. Furthermore,
233 ΔI values are close to 0 for durations longer than about 300 ms whereas they vary from -40 to

234 40 dB for shorter durations. Actually, standard deviations of ΔI values (Figure 2, bottom) are
235 small above 300 ms, and larger under 300 ms. The shift from small to large ΔI values is stable
236 around 300 ms for all CVCV sequences, leading us to select this value as a limit for further
237 analyses. Thus, for slow rates, the speakers keep a regular alternation of labial and coronal
238 syllables without any strong reduction effect. Then, rate increase induces the reduction of either
239 V_{La} (positive ΔI values) or V_{Co} (negative ΔI values). Hence, both the L_aC_o (/psá/) and the C_oL_a
240 (/spá/) CCV clusters seem to constitute possible attractors for the repetition of either L_aC_o or C_oL_a
241 CVCV logatoms. Yet, the graphs display more positive than negative ΔI values for the two types
242 of disyllables. Therefore, the L_aC_o attractor seems to prevail over the C_oL_a one, particularly for
243 the L_aC_o trials (Figure 2, top left).

244 ---- Figure 2 ----

245

246 2.2.2. Detailed analysis according to the disyllables and to the speakers

247 In order to investigate more precisely possible reorganizations of productions in the
248 speeding process, the analysis of ΔI was restricted to utterances with durations shorter than
249 300 ms. In spite of the flash, the utterance durations varied a lot according to the speaker and to
250 the trial. Therefore, for the analysis, we kept only speakers achieving a sufficient amount of quick
251 data with the following criterion: for each of the six CVCV logatoms, the speaker should have
252 produced at least five quick utterances (durations shorter than 300 ms) for at least two of his/her
253 three trials. Then, only the two trials with the largest number of quick utterances were analyzed.
254 This criterion excluded five speakers. In the following, only the utterances with durations less
255 than 300 ms are analyzed, for the two selected trials of the 21 remaining subjects.

256 Global means of the 42 trials are positive for all disyllables, and all significantly different
257 from 0 except for /sapa/ (Table I, first row). Moreover, the distribution of the 42 means according
258 to their sign (last four rows in Table I) shows that the ratio of positive means related to negative
259 ones (rows 3 and 5) is 2.5 for /pata/ and /tapa/, 7.4 for /fata/, 3.6 for /tafa/ and 1.8 for /pasa/
260 whereas it is 1 for /sapa/. A Chi-square test shows significant difference between the frequencies
261 of positive vs. negative means ($\chi^2(1) = 6.6, p < .001$). Restricted to means that significantly differ
262 from 0 (rows 4 and 6) the prevalence of positive means over negative ones is even greater with a
263 ratio of 4.75 for /pata/, 2.5 for /tapa/, 3 for /pasa/, 1.4 for /sapa/, 13 for /fata/ and 5.3 for /tafa/
264 (frequency of significant positive means significantly different from significant negative ones,
265 $\chi^2(1) = 6.8, p < .001$). Hence, the speeded repetition of L_aC_o and C_oL_a CVCV disyllables more
266 often evolves towards the corresponding L_aC_o attractor (eg. /pasá/ or /psá/) than the C_oL_a one (eg.
267 /sapá/ or /spá/). However, C_oL_a forms appear in some cases for both L_aC_o and C_oL_a trials.
268 Furthermore, two additional factors tend to emerge. Indeed, there is some trend that ΔI values are
269 larger, and more systematically positive, for L_aC_o than for C_oL_a trials on one hand, and for /pata/-
270 /tapa/ and /fata/-/tafa/ than for /pasa/-/sapa/ trials on the other hand.

271

272 ---- Table I ----

273

274 Considering that two attractors seem to coexist, though with a preference for the L_aC_o
275 one, we wondered whether they might coexist for a given speaker. For this aim, Figure 3 (left)
276 displays the relation between mean ΔI values for the first and the second selected trial, for each
277 speaker and for each CVCV sequence, provided that the two means significantly differed from 0.

278 It appears that mean ΔI values are both significantly positive for 41 cases, both significantly
279 negative for 7 cases and have different signs in 5 cases. Therefore, the dominant behavior is to
280 evolve towards the same attractor for the two trials with a large preference for the L_aC_o compared
281 to the C_oL_a one. Yet, in some cases, the attractor differs from one trial to the other. Then,
282 grouping all logatoms, it appears that some speakers tend to be “ L_aC_o stable” (e.g. S12, right part
283 of the Figure 3) whereas others are, in a lower extend, rather “ C_oL_a stable” (e.g. S23). Finally,
284 some speakers are “bi-stable”: the category of attractor they evolved towards changes according
285 to the trial (S20) or to the logatom (S6).

286

287 ----- Figure 3 -----

288

289 In summary, the speeded repetition of L_aC_o and C_oL_a logatoms could evolve both
290 towards the L_aC_o or the C_oL_a attractor. Yet, and despite variability between disyllables and
291 between speakers, the L_aC_o attractors are largely favored compared to the C_oL_a ones. Thus, /patá/-
292 /ptá/ is the favorite attractor for /pata/ and /tapa/; /pasá/-/psá/ for /pasa/ and to a lower extend for
293 /sapa/ and /fatá/-/ftá/ for both /fata/ and /tafa/.

294 3. An articulatory study of L_aC_o and C_oL_a stability

295 The main purpose of the second study is to understand from EMMA recordings how the
296 articulatory system could evolve with rate increase from a L_aC_o or a C_oL_a CVCV disyllable
297 towards either a L_aC_o or a C_oL_a attractor, with a preference for the first one.

298 3.1. Method

299 3.1.1. *Speakers and phonetic material*

300 The participants were five college-aged French speakers (three females - S1, S3 and S4 -
301 and two males - S2 and S5), without any speech or hearing deficits. They did not participate in
302 Experiment 1. The L_aC_o and C_oL_a disyllables were the same as in the acoustic study. Two
303 duplicated CVCV logatoms (/papa/ and /tata/) were added as control items.

304 3.1.2. *Procedure, data recording and processing*

305 The instruction was to continuously repeat the disyllable enounced by the experimenter
306 starting at a slow rate and then at an increasingly rapid rate up to the highest possible one. The
307 speakers then had to progressively decrease the rate to return to the initial slow one. Contrary to
308 the first experiment, no external tempo driver was used because of the already difficult conditions
309 of EMMA recording. Speakers were encouraged not to stop their productions even if they seemed
310 different from the initial sequence. The eight disyllables were repeated three times with orders
311 that varied from one subject to another. The articulator motions were tracked over time at 500 Hz
312 using a Cartens Electromagnetic Articulograph. Two reference transducers, one at the bridge of
313 the nose and the other on the gums above the upper teeth allowed to correct the data for head
314 movements. The three measurement points were: the jaw (transducer placed beneath the lower
315 teeth); the Tongue Tip (TT, transducer placed at about 1 cm from the tip of the tongue) and the
316 Lower Lip (LL, transducer just beneath the vermilion border of the lower lip). These raw
317 measurements were then processed according to classical procedures (see Hoole P., 1996 for
318 more details). These involved translation and rotation correcting for head movement, resulting in

319 setting the origin at the reference coil located on the upper incisors, the vertical axis as the line
320 joining the reference coils on upper incisors and nose bridge, and the x-axis aligned with the
321 occlusal plane. The acoustic signal was simultaneously recorded by the way of a microphone
322 fixed on the articulograph helmet and then digitized at 20 kHz.

323 3.1.3. *Hypotheses*

324 The first hypothesis is that the mandible would evolve from a one-cycle-per-syllable
325 relationship to a one-cycle-per-disyllable relationship when speeding L_aC_o or C_oL_a CVCV
326 logatoms. However, duplicated L_aL_a (/papa/) or C_oC_o (/tata/) CVCV logatoms were expected not
327 to display this behavior. The second hypothesis concerns the lips-tongue coordination inside a
328 single jaw cycle at high rates. Two attractors were expected, either L_aC_o or C_oL_a . In the first case,
329 the labial constriction and the next coronal constriction (respectively maximum values of LL and
330 TT trajectories) should be close together. In the second case, the coronal constriction should be
331 close to the next labial one. Considering the small number of speakers, and the variability
332 observed in the first experiment, no strong claim was made about a preference for L_aC_o attractors.
333 Instead, the hope was that inter-speaker variability would enable to observe both kinds of
334 attractors, in order to be able to characterize them articulatorily.

335

336 ---- Figure 4 ----

337

338 *3.1.4. Measurements*

339 Processing of acoustical data was the same as in Experiment 1. It led to computation of
340 the duration and intensity variation between the two vowels (ΔI) for each utterance (CVCV or
341 CCV items correctly produced). Articulatory analysis used articulator trajectories preprocessed
342 by a low-pass Chebychev filter (cutting frequency at 15 Hz). Jaw maxima and minima (Figure 4,
343 row 2) were automatically detected on the jaw trajectory, allowing its segmentation into closing
344 and opening phases (up and down strokes). Post-processing discarded movements realized during
345 pauses or errors (detected from the acoustical labeling). For the correct utterances, two criteria
346 were used in order to discard productions subtended by too small jaw movements. Firstly, we
347 checked for the regular alternation of opening and closing strokes; secondly, we removed strokes
348 with an amplitude lower than 0.1 mm, which is the claimed precision of the EMMA system.
349 Table II summarizes the total number of correct utterances and the number of correct utterances
350 kept for further articulatory analysis. Then, TT and LL events were labeling on articulatory
351 signals, with the help of acoustic events. TT constrictions were defined as the highest TT
352 positions between the vowel after the labial consonant (V_{La}) and the vowel after the coronal
353 consonant (V_{Co}) (Figure 4, row 3). Similarly, LL constrictions were defined as the highest LL
354 positions between V_{Co} and V_{La} (Figure 4, row 4). This method (similar to Hertrich and
355 Ackermann's one, 2000) discards secondary maxima due to passive motion induced by the jaw
356 gesture that comes with the other constriction.

357

358 ---- Table II ----

359

360 Then, jaw cycles were separated into an up stroke and its adjacent down stroke. The jaw
361 cycle duration was the sum of its up and down stroke durations and the jaw cycle amplitude was
362 the mean of its up and down stroke amplitudes (Figure 5.A). Each constriction event was
363 assigned to the jaw cycle that temporally included it and positioned relatively to this jaw cycle
364 (Figure 5B). This measure, expressed in percent of jaw cycle, constitutes an indicator of the
365 phasing relationship between the jaw and constrictors. Finally, the coordination between TT and
366 LL was characterized by the duration of the labial-to-coronal phase (time from LL constriction to
367 TT constriction D_{LaCo}), normalized by the whole disyllable duration (time from a LL constriction
368 to the next LL one D_{LaLa}):

$$369 \quad \Delta_{(LaCo/LaLa)} = D_{LaCo} / D_{LaLa}$$

370

371 ----- Figure 5 -----

372

373 3.2. Results

374 3.2.1. Acoustic analysis

375 The global pattern of ΔI against utterance duration for L_aC_o or C_oL_a sequences is very
376 close to the pattern observed in the first experiment (compare Figure 6, right, with Figure 2). That
377 is, ΔI values are close to 0 for durations longer than 300 ms whereas ΔI values far from 0 appear
378 for shorter durations, with more positive than negative values. On the contrary, ΔI values are
379 always close to 0 for duplicated L_aL_a or C_oC_o logatoms (Figure 6, left).

380

381

---- Figure 6 ----

382

383

384

385

386

387

388

389

390

391

392

---- Table III ----

393

394

3.2.2. *Global analysis of jaw cycles*

395

396

397

398

399

400

401

Moreover, the mean values for utterances with durations shorter than 300 ms (Table III) represent the panel of behaviors observed in the first experiment. Indeed, as a subset of speakers of the acoustic study, S3 fails to produce enough rapid productions. This may indicate that this speaker rarely shifts towards a CCV structure. Then, S1 and S4 tend to favor the C_oL_a attractor (more negative than positive means) especially for /pasa/ and /sapa/ for S1 and for the three C_oL_a disyllables for S4. On the contrary S2 and S5 tend to favor the L_aC_o attractor for all disyllables except for S5's /sapa/ productions. Interestingly, these behaviors correspond to both the speakers' profiles and the variability between sequences observed in the first experiment.

The /tapa/ utterance samples displayed on Figure 4 illustrate the speeding process generally observed for a L_aC_o or a C_oL_a trial. First, during the acceleration phase (left), each constriction comes with its own jaw gesture. Yet, the amplitude and the duration of the jaw cycle for the labial constrictions tend to be smaller than for the coronal constrictions. Then, rate increase induces the progressive fading of the jaw cycle that supports the labial constriction, resulting in a single jaw cycle for the two constrictions. For this example, the plot of the acoustic signal (top) shows that the speaker does not evolve towards a CCV syllable but produces two

402 syllables on a single jaw cycle (splitting phase, middle). Finally, the deceleration progressively
403 brings the system back to its initial state with one jaw cycle per syllable (right).

404 Plots of jaw cycle durations (y) against CVCV durations (x) for all speakers' productions
405 (Figure 7, first row) and separately for the duplicated (left) and variegated (right) CVCV
406 sequences show that duplicated productions are grouped around the ($y=x/2$) line (i.e. two jaw
407 cycles for one disyllable) whereas both L_aC_o and C_oL_a ones are distributed around the ($y=x/2$) and
408 ($y=x$) lines (i.e. two vs. one jaw cycle per disyllable). This distribution splitting occurs for
409 utterance durations less than about 400-450 ms. Analyses of ΔI in the acoustic study showed that
410 the shift towards a CCV cluster, characterized by high ΔI values, occurs for durations shorter than
411 300 ms. Therefore, the shift towards a single cycle for two syllables occurs before the total
412 suppression of one of the two vowels (that is, for longer utterance durations).

413

414 ----- Figure 7 -----

415

416 The utterance duration histograms (Figure 7, second row) display similar values for the
417 duplicated and the L_aC_o / C_oL_a group. Yet, the durations of duplicated utterances are never shorter
418 than 225 ms and are mostly around 300 ms whereas around 15 % of L_aC_o / C_oL_a productions
419 reach durations from 175 ms to 225 ms with a peak of distribution around 250-275 ms. On the
420 contrary, the jaw cycle durations (Figure 7, third row) are clearly shorter for the duplicated
421 productions (peak of distribution around 150 ms) than for L_aC_o / C_oL_a ones (peak of distribution
422 around 225 ms). Finally, the jaw cycle duration is never shorter than 100 ms for all logatoms.
423 Altogether, this portrait is coherent with the fact that at a fast rate L_aC_o / C_oL_a sequences can be

424 produced on one single jaw cycle. The progression towards a single cycle limits jaw motion and
425 keeps it at a more comfortable duration for the fastest production rates. Thus, it probably limits
426 the energy consumption.

427 In order to analyze the shift from two jaw cycles to a single one, the speaker's productions
428 were classified into three periods (e.g. Figure 4). The "splitting" phase includes the utterances
429 realized on a single jaw cycle. The "acceleration" and the "deceleration" phases include the
430 utterances realized on two jaw cycles respectively before and after the splitting phase.

431 3.2.3. *Patterns of jaw motion in the acceleration phase*

432 The amplitude and duration of jaw opening and closing strokes in the acceleration phase
433 are plotted on Figure 8, for each speaker, respectively for duplicated (left) and L_aC_o / C_oL_a (right)
434 sequences. Strokes are seldom shorter than 50 ms, which agrees with the 100 ms threshold
435 observed for the jaw cycle durations (Figure 7). Separate regression lines are superimposed for
436 labial and coronal cycles. In their study of jaw motion for duplicated sequences (/sa/), Nelson *et*
437 *al.* (1984) used the same kind of display. A theoretical analysis allowed the authors to define a
438 limit parabolic curve, such that for a given value of the maximum acceleration during a jaw
439 stroke, the corresponding amplitude vs. duration points were always on the right of the limit
440 parabolic curve. Without entering their explanation in too many details, the principle is basically
441 that to achieve a given trajectory extent in a given duration, the acceleration along the trajectory
442 must achieve a sufficient maximum value. Reciprocally, for a given maximum acceleration value
443 and a given trajectory extent, duration cannot be lower than a limit, specified by the parabolic
444 curve. On the figure, we have superimposed three such limit curves, respectively corresponding
445 to maximum acceleration values of 0.25 g, 0.50 g and 1.5 g (see also Nelson, 1983).

446 It appears that for duplicated utterances (Figure 8, left), speakers S3 and S4 follow a
447 different pattern from speakers S1, S2 and S5. For S3 and S4, stroke amplitude decreases with
448 duration for both the labials (in /papa/) and the coronals (in /tata/) cycles. Moreover, the
449 distribution of amplitude vs. duration tends to follow the theoretical curve corresponding to an
450 acceleration of 0.25 g. On the contrary, amplitude does not decrease with duration for S1, S2 and
451 S5. Moreover, while S1 keeps a small amplitude for all durations, rate increase leads S2 and S5
452 beyond the 0.25 g limit curve, towards an acceleration up to 0.5 g. Therefore, to support rate
453 increase, speakers either decrease the jaw stroke amplitude without increasing acceleration (S3
454 and S4) or increase acceleration without changing amplitude (S1, S2 and S5). These results agree
455 with the study by Sonoda and Nakakido (1986). These authors observed that jaw motion speeding
456 induced either an increase in velocity and little change in amplitude or little velocity change and a
457 decrease in amplitude (see also Lindblom, 1990a for a link with phonetic variations; and Perkell
458 et al. 2002 for a more detailed discussion of intersubject differences).

459

460 ---- Figure 8 ----

461

462 The jaw behavior is more homogeneous for the L_aC_o / C_oL_a group. For all speakers, there
463 is a global decrease of amplitude with duration (Figure 8, right) and the distributions more or less
464 follow the 0.25 g curve. Furthermore, the means of amplitude and duration tend to be smaller for
465 the strokes that subtend the labial constrictions as compared to the strokes that subtend the
466 coronal constrictions, with significantly negative differences in a number of cases, and no
467 significantly positive difference (Table IV). This pattern is modulated by two factors. Firstly,

468 there are individual differences, with, for example, larger differences between the labial and
469 coronal jaw strokes for speaker S3, and smaller for speaker S2. Secondly, the labial to coronal
470 difference is increased for /pasa/ and /sapa/ logatoms (for which the labial jaw cycle is of a
471 significantly lower amplitude and duration for all speakers) and decreased for /fata/ and /tafa/
472 logatoms (for which the difference is significant only for Speaker 3). This portrait suggests that in
473 the acceleration phase, the passage from two jaw cycles to a single one rather corresponds to a
474 progressive removal of the “labial cycle”, as displayed in Figure 4. We shall now explore how the
475 reorganization between jaw, tongue and lips occurs in the “splitting” phase.

476

477 ----- Table IV -----

478

479 *3.2.4. Coordination between jaw and constrictors in the acceleration, splitting and*
480 *deceleration phases*

481 For most of the productions realized on two jaw cycles (during the acceleration and the
482 deceleration phases), the TT and LL constrictions are around 0% of the jaw trajectory (Figure 9,
483 top and bottom). This indicates that each constriction occurs when the mandible is at its highest
484 position. So, TT and LL successively move in phase with the jaw. Yet, distributions look less
485 peaked at 0% for TT (Figure 9, right top and bottom) than for LL (Figure 9, left top and bottom)
486 for both duplicated and L_aC_o / C_oL_a groups. Moreover, for L_aC_o and C_oL_a logatoms, there was
487 some trend for more variations during the deceleration phase than during the acceleration phase.
488 This could be interpreted as a sign of hysteresis, considering the much more complex portrait in
489 the splitting phase.

490 In the splitting phase, the labial and coronal constrictions must reorganize within one jaw
491 cycle. The basic trend for reorganization is clearly displayed in Figure 9: the tongue constriction
492 is around the highest jaw position (phase 0%), while the labial constriction is around the lowest
493 position (phase 50% or -50%). On this basic pattern, a global shift towards the right of the jaw
494 cycle can appear, with a coronal constriction in the beginning of the opening phase (around 25%)
495 and a labial constriction in the beginning of the closing phase (around -25%). Two factors seem
496 to intervene in this possible shift. The first factor is speaker variability, with S2 more involved in
497 the right shift than all other speakers. The second factor is the nature of the labial and coronal
498 consonants. Thus, for /pasa/ and /sapa/, the coronal constriction is strictly located around the jaw
499 highest position (0% phase) while for /fata/ and /tafa/ it is often shifted towards positive values.
500 The correspondence with the predominance of the coronal jaw cycle upon the labial one in terms
501 of amplitude and duration, displayed in the previous section, is obvious. It is rather
502 straightforward. Indeed, the coronal fricative /s/ induces a specifically high jaw position,
503 generally higher than all other phonemes, including the coronal plosive /t/ or /d/ (Keating, 1983;
504 Lindblom, 1983; Keating *et al.*, 1994; Lindblad and Lundquist, 1999, see also data for French in
505 Rhardisse and Abry, 1994). Hence the coronal jaw cycle is increased, and the coronal constriction
506 is more strictly stuck on the highest jaw value. On the contrary, the labiodental /f/, also inducing
507 a high jaw position, increases the labial jaw cycle, which induces the shift towards the right, with
508 higher values of the labial constriction in the jaw trajectory. This global pattern is also modulated
509 by the difference between L_aC_o and C_oL_a attractors.

510

511

---- Figure 9 ----

512

513 *3.2.5. Articulatory characterization of L_aC_o and C_oL_a attractors*

514 L_aC_o utterances should be characterized by both a higher intensity of the second vowel
515 over the first one (positive value of ΔI , Table III) and a proximity of the labial and next coronal
516 constrictions ($\Delta_{(L_aC_o/L_aL_a)}$ smaller than 50%, Table V). Notice however that data selection for
517 these two analyses satisfy different criteria, that is utterance duration less than 300 ms for the
518 acoustic analysis and utterances produced on a single jaw cycle for the articulatory analysis.
519 Despite these differences, the two analyses are relatively coherent.

520 Indeed, means of $\Delta_{(L_aC_o/L_aL_a)}$ tend to be higher than 50% for S1, which agrees with the
521 C_oL_a profile that characterizes this speaker in the acoustic space. On the contrary, all $\Delta_{(L_aC_o/L_aL_a)}$
522 means for S2 are smaller than 50%, which confirms the speaker's tendency to be L_aC_o -stable. For
523 S4, means of $\Delta_{(L_aC_o/L_aL_a)}$ are smaller than 50% for all L_aC_o sequences whereas they are greater
524 than 50% for /sapa/ and /tafa/, which agrees with acoustic data showing that this speaker tends to
525 be L_aC_o -stable for the L_aC_o logatoms and C_oL_a -stable for the C_oL_a ones. For S5, $\Delta_{(L_aC_o/L_aL_a)}$
526 means are smaller than 50% for all the sequences except /pasa/ and /sapa/, which confirms his
527 acoustic profile: L_aC_o for /pata/, /tapa/, /fata/ and /tafa/ but rather C_oL_a for /sapa/. Finally, results
528 for S3 show that though this speaker does not often reach utterance durations shorter than 300 ms
529 (Table III), she evolves towards a single jaw cycle for all sequences. Moreover, she tends to be
530 L_aC_o -stable with $\Delta_{(L_aC_o/L_aL_a)}$ values always significantly smaller than 50% except for /fata/.

531

532

---- Table V ----

533
534 Therefore, quite logically the L_aC_o attractor corresponds to a decrease of the temporal
535 distance between the labial and the next coronal constriction, while the C_oL_a attractor corresponds
536 to a decrease of the temporal distance between the coronal and the next labial constriction.
537 However, there is a clear predominance of switches towards the L_aC_o attractor (e.g. /patá/ or /ptá/)
538 over the 21 speakers of the acoustic study (Experiment 1). This agrees with the hypothesis
539 according to which the anticipation of the tongue gesture during the labial constriction, typical of
540 L_aC_o CVCV' (e.g. /patá) or CCV' (e.g. /ptá) sequences, is much more likely than the anticipation
541 of the labial gesture during the tongue constriction.

542 4. Discussion

543 The present work aimed at testing the hypothesis that the L_aC_o sequencing is a more
544 stable articulatory pattern than the C_oL_a one. The results of Experiment 1 provide arguments for
545 this hypothesis for French. Furthermore, the results of Experiment 2 allow to better describe the
546 articulatory phenomena that subtend this asymmetry.

547 4.1. The L_aC_o / C_oL_a coordinative system “on line”

548 The results of the articulatory study confirm that rate increase could induce the repetition
549 of L_aC_o and C_oL_a CVCV sequences to evolve from one jaw cycle per syllable towards a single
550 jaw cycle for two syllables until a single CCV syllable is eventually achieved. This progression
551 seems more economic regarding jaw motion. In this process, the determination of the attractor
552 (L_aC_o vs. C_oL_a) may result from the interaction of several factors.

553 4.1.1. *The motor program shift*

554 At the beginning of the repetition process, the speakers produce one jaw cycle per
555 syllable. During this period, each constriction occurs at the highest jaw position. Then, with rate
556 increase, the amplitude and the duration of the jaw strokes decrease together until reaching a
557 threshold value (near 50 ms for an opening or a closing gesture). Coupled with the anticipation of
558 one constriction during the previous one, this induces the shift on a single cycle. This shift rests
559 on the reorganization of the tongue-lip-jaw coordination. Globally, two coordinative patterns are
560 displayed to achieve the two constrictions on a single jaw cycle (Figure 10, right), depending on
561 the initial profile of jaw motions (Figure 10, left). Firstly, and in most of the cases, the jaw cycle
562 that carries the labial constriction becomes shorter and smaller than the jaw cycle that carries the
563 coronal constriction (Figure 10, top left). Hence, when rate increases, the labial cycle reaches the
564 threshold duration earlier than the coronal one. This induces the preservation of the phasing
565 between the jaw and the tongue tip highest positions, whereas the lower lip constriction occurs
566 during the jaw closing gesture or at the end of the opening gesture (Figure 10, top right). This is
567 typically the case for Speaker 3, and for /pasa/ and /sapa/ utterances. Secondly, when no
568 asymmetry appears between the two cycles during the acceleration phase (Figure 10, bottom left),
569 the tendency is to observe a dephasing of both TT and LL in reference to the jaw cycle. Thus, the
570 LL and TT constrictions respectively occur during the closing gesture and during the opening
571 gesture (Figure 10, bottom right). This is the trend for Speaker S2 and for /fata/ and /tafa/
572 utterances.

573

574 ---- Figure 10 ----

575
576 These results are interesting in the context of previous articulatory studies (Nelson *et*
577 *al.*, 1984; Kelso *et al.*, 1985; Hertrich and Ackermann, 2000), modeling works (Redford and
578 Diehl, 1999; Redford *et al.*, 2001; Oudeyer, 2001) or theories (MacNeilage, 1998) that always
579 associated the jaw cycle to the CV syllable. Moreover, patterns of mandible motions are close to
580 data published in the literature and tend to confirm the general principle of economy of effort.

581 4.1.2. *Economy of effort*

582 Various possible indicators of economy of effort have been proposed in the literature, e.g.
583 maximum speed or jerk that is acceleration derivative (Nelson, 1983). In their already mentioned
584 study, Nelson *et al.* (1984) characterized jaw strokes by the maximum acceleration compatible
585 with the observed movement, which is the indicator we used in Section 3. Furthermore, they
586 showed that during the repetition of /sa/ with rate increase, the jaw motion progressed in a way
587 that tended to minimize energy consumption. In their data, the durations of jaw opening and
588 closing strokes were never shorter than 50 ms. Similar values appear in the present articulatory
589 study. Then, for one of their three analyzed speakers, Nelson *et al.* observed “a maximum in the
590 movement distance as the movement times decreased into the region of 125 ms”. According to
591 the authors, this “resonance” effect shows that 125 ms could be the most economic duration for
592 up and down jaw strokes. It would correspond to the “natural” rate of 3-4 cycles and so, 3-4
593 syllables per second. In our study (Experiment 2), durations of jaw cycles for L_aC_o and C_oL_a
594 utterances are mainly around 225 ms (Figure 7), which corresponds to jaw stroke durations
595 around 112.5 ms, close to the 125 ms value observed by Nelson *et al.* On the other hand, the
596 production of duplicated CVCV sequences at very fast rates requires to decrease the jaw cycle

597 down to 150 ms and so, to stroke durations around 75 ms, close to the saturation value.
598 Furthermore, estimation of the maximal acceleration value in Figure 8 indicates that it is
599 generally smaller for the L_aC_o and C_oL_a sequences compared to the duplicated ones. Taken
600 together, these results suggest that when rate increases, the coordination of the two constrictions
601 on a single jaw cycle might be more economic for the jaw than the alternation of two jaw cycles.

602 *4.1.3. Multi-determinism of a bistable system*

603 The results show that with rate increase, the lips-tongue-jaw system could evolve either
604 towards a L_aC_o attractor (e.g. /pata/→/patá/→/ptá/) or a C_oL_a one (e.g. /pata/→/tapá/→/tpá/).
605 Thus, the coordinative system is bistable. Yet, according to the first study, the L_aC_o attractor
606 prevails. On this basic pattern, a number of other additional factors are observed or could be
607 assumed to play a role. Firstly, the initial coordination provides a bootstrap towards the
608 corresponding attractor: the preference for the L_aC_o attractor is less marked for the repetition of
609 C_oL_a CVCV sequences as compared to L_aC_o ones. This is particularly the case for /pasa/ and
610 /sapa/, with /pasa/ more frequently evolving towards the L_aC_o attractor whereas /sapa/ often
611 evolved towards the C_oL_a one. This is likely to be due to the fact that, in spite of the instruction to
612 continue even if transformations occurred, some speakers tried to respect the initial structure and
613 even tended to reduce the speed when they perceived a shift to the reverse sequencing. Secondly,
614 the C_oL_a attractor was more often observed for /sapa/ as compared to /tapa/ and /tafa/. The
615 difference between /sapa/ and the other C_oL_a disyllables could rest on articulatory factors. Indeed,
616 analysis of the jaw cycles in the second experiment showed that in the acceleration phase, the
617 asymmetry between the jaw cycles that come with the labial constriction and the jaw cycles that
618 come with the coronal one are more systematic for /pasa/-/sapa/ than for the other sequences,

619 which was interpreted in terms of the need for a specific high jaw position for achieving /s/. This
620 could favor the progression towards the C_oL_a attractor for /pasa/ and /sapa/. Thirdly, inter-speaker
621 differences, which could come from specific individual patterns of coordination, lead to
622 differences in the selection of one or the other attractor, as displayed in Section 2.2.2.

623 But a number of other factors, not studied in the present work, could also intervene in the
624 selection of one or the other attractor. An obvious one is the selected language. Indeed, it is likely
625 that the CCV re-syllabification process observed here results from the fact that these forms are
626 possible in French and that the language stresses the last syllable of a word. Therefore, future
627 experiments should be done to assess the L_aC_o greater stability in languages with other prosodic
628 behaviors or phonological constraints. Other additional experiments should also be done on the
629 role of the vowel context, the consonantal mode (e.g. voiced or nasal consonants), etc.
630 Nevertheless, in spite of all these established or potential additional factors, the L_aC_o attractor is
631 globally the dominant one over all speakers and disyllables in the present study, which suggests
632 that it could indeed be the most stable and economic, at least in French.

633 4.2. Substance-based selection of phonological sequences

634 The underlying assumption of this work is that world languages inventories might be
635 shaped by substance-based constraints such as properties of the motor control system. Thus, some
636 sequencing of syllables might be preferred because they are easier to produce in the sense that
637 they are more economic (less energy consuming) or more stable (easier to control). In this
638 framework (Liljencrants and Lindblom, 1972; Lindblom *et al.*, 1984; Kelso., 1995; Davis *et al.*
639 2002), the present study brings at least two interesting contributions: (1) the production of
640 duplicated CVCV sequences may be more costly than the production of CVCV sequences that

641 vary the constrictor from the first to the second consonant, and (2) the L_aC_o sequencing seems
642 more stable than the reverse C_oL_a one.

643 *4.2.1. Variegated rather than duplicated*

644 Statistical analyzes of world languages lexical inventories show that in C_1VC_2V
645 sequences, lexicons favor associations that vary the articulation places in C_1 and C_2 , rather than
646 those that repeat the same constriction (MacNeilage and Davis, 2000a, 2000b; Rousset 2003).
647 This means that languages prefer e.g. “pata” to “papa”. Yet, duplication could, a priori, be
648 considered as simpler in terms of control, which could explain that it is the core behavior in
649 canonical babbling (MacNeilage and Davis, 2000a, 2000b). However, the articulatory data
650 obtained here lead to propose that the preference for variegated sequencing in world languages
651 could result in part from the fact that the repetition of the same constriction is more costly for the
652 jaw than the alternation of two different constrictions. Indeed, the comparison of the production
653 of L_aC_o / C_oL_a and duplicated CVCV sequences shows that quicker production rates could be
654 achieved for L_aC_o / C_oL_a sequences with less energy consumption for the jaw, since anticipation
655 allows producing L_aC_o and C_oL_a CVCV utterances on a single jaw cycle whereas it is impossible
656 for duplicated ones. Thus, without involving the extreme speeds obtained here at the end of the
657 acceleration phase, it appears that the alternation of the movement of two constrictors is probably
658 easier and more economic than the repetition of the movement of a single constrictor for the
659 mature speech production system. Similar principles govern other synergies such as fingers in
660 piano playing. This results from the fact that when two different organs are involved, the motion
661 of one organ could be anticipated during the realization of the other one, which is not possible
662 when a single organ realizes the two successive gestures. The possibility to produce two syllables

663 on a single jaw cycle might be generalized to other kinds of sequencing than L_aC_o and C_oL_a ones
664 in further investigations. Moreover, the preference for the L_aC_o attractor in the two experiments
665 shows that inside the preference for variegated forms, some associations might be more
666 appropriate than others.

667 *4.2.2. L_aC_o rather than C_oL_a*

668 The results of Experiment 1 are in favor of the idea that L_aC_o sequences appear as a
669 natural and coherent production unit, more stable and better “in-phase” than their C_oL_a inverse
670 counterpart. This could explain the preference for L_aC_o utterances in infant production, rather
671 than the “simple first” explanation previously proposed. There is actually no reason to believe
672 that the greater stability of L_aC_o sequences in the present experiments on adults could be due to
673 any “simple first” mechanism. Rather, articulatory coherence could provide a common basis to
674 both infant data and our results. Of course, the fact that the C_oL_a structure may also be locally
675 stable, though less often selected in the speeding process, is not incompatible with this reasoning.
676 It just suggests that C_oL_a sequences are also viable, though suboptimal in comparison with L_aC_o
677 ones, which agrees well with typological and developmental data.

678 5. Conclusion

679 The experimental data analyzed here point out a major difference between L_aC_o and C_oL_a
680 sequences at the production level: L_aC_o sequences can be considered as a more stable
681 coordinative pattern for the motor system than C_oL_a ones. Then, a causal relationship between
682 this asymmetry and the universal “LC effect” has been advanced. Of course, such a claim has to
683 be improved testing L_aC_o and C_oL_a sequences stability in both other languages and more directly

684 in infant' first words. To this aim, the paradigm and the acoustic analysis method proposed here
685 could allow comparing L_aC_o and C_oL_a stability for speaker samples of different languages, which
686 is necessary for generalization. Specific methodologies should be developed for studying stability
687 in infant speech. Moreover, further articulatory investigations of jaw-tongue-lip coordination in
688 the production of C_1VC_2V disyllables would be necessary to better understand the anticipation
689 processes, and their link with the asymmetry in jaw cycles for different constriction places in C_1
690 and C_2 . It remains that the present study adds some new material to the already significant set of
691 perceptuo-motor constraints likely to play a role in the shaping of phonological systems,
692 hopefully taking substance-based approaches to language universals one step further.

693 6. Acknowledgements

694 This work is part of the “Patipapa” project funded by the French Ministry of Research
695 (Action Concertée Incitative “Systèmes Complexes en Sciences Humaines et Sociales”). We also
696 thank C. Savariaux and C. Vilain for their technical support and all the speakers for their essential
697 participation.

698

699

699 7. Notes

700 ¹The Praat software is developed by P. Boersma and D. Weenink, at the Institute of
701 Phonetic Sciences. It is available to free download on their web page
702 (<http://www.fon.hum.uva.nl/praat/>).
703

703

704 8. References

705 Byrd, D. (1996). “Influences on articulatory timing in consonant sequences,” J.
706 Phonetics, **24**, 209-244.

707 Chitoran, I., Goldstein, L. and Byrd, D. (2002). *Gestural overlap and recoverability:*
708 *Atriculatory evidence from Georgian*, in C. Gussenhoven & Warner (Eds),
709 Laboratory Phonology 7. Berlin: Mouton de Gruyter, 419-448.

710 Clements, G. N. (1990). “The role of the sonority cycle in core syllabication,” in *Papers*
711 *in Laboratory Phonology I: Between the Grammar and the Physics of Speech*, edited
712 by Kingston, J. & Beckman, M., (Cambridge University Press), 283-333.

713 Davis, B.L., MacNeilage, P.F., and Matyear, C. (2002). “Acquisition of serial complexity
714 in speech production: A comparison of phonetic and phonological approaches to first
715 word production,” *Phonetica*, **59**, 75-107.

716 de Jong, K. J. (2001). “Effects of syllable affiliation and consonant voicing on temporal
717 adjustment in a repetitive speech-production task,” *J. Speech Lang. Hear. Res.* **44**,
718 826-840.

719 Haken, H., Kelso, J. A. S. and Bunz, H. (1985). “A theoretical model of phase transitions
720 in human hand movements,” *Biol. Cybern.* **51**, 347-356.

721 Hertrich, I. and Ackermann, H. (2000). “Lip-jaw and tongue-jaw coordination during
722 rate-controlled syllable repetitions,” *J. Acoust. Soc. Am.* **107**, 2236-2247.

723 Hoole P. (1996) "Issues in the acquisition, processing, reduction and parameterization of
724 articulographic data," FIPKM 34, pp 158-173

725 Hoyt, D. and Taylor, C.R. (1981). "Gait and the Energetic of Locomotion in Horses,"
726 Nature, **292**, 239-240.

727 Ingram, D. (1974). "Fronting in child phonology," J. Child Lang. **1**, 233-242.

728 Keating, P., Lindblom, B., Lubker, J., and Kreiman, J. (1994). "Variability in jaw height
729 for segments in english and swedish VCVs," J. Phonetic, **22**, 407-422.

730 Keating, P. (1983). "Comments on the jaw and syllable structure," J. Phonetic, **11**, 401-
731 406.

732 Kelso., J. A. S. (1995). *Dynamic patterns: The self-organization of brain and behavior*.
733 (MIT Press), Cambridge, MA.

734 Kelso, J. A. S., Saltzman, E. L. and Tuller, B. (1986). "The dynamical perspective on
735 speech production: data and theory," J. Phonetic, **14**, 29-59.

736 Kelso, J. A. S., Vatikiotis-Bateson, E., Saltzman, E. L. and Kay, B. (1985). "A qualitative
737 dynamic analysis of reiterant speech production: phase portraits, kinematics, and
738 dynamic modeling," J. Acoust. Soc. Am. **77**, 266-280.

739 Liljencrants, J. and Lindblom, B. (1972). "Numerical simulations of vowel quality
740 systems: the role of perceptual contrast," Language, **48**, 839-862.

741 Lindblad, P. and Lundqvist, S. (1999). "How and why do the tongue gestures of [t], [d],
742 [l], [n], [s], and [r] differ?", in *Proceedings of the 14th ICPhS*, 417-420.

743 Lindblom, B. (1990a). "Explaining Phonetic Variation, A Sketch of the H&H Theory," in
744 *Speech Production and speech modeling*, edited by W. J. Hardcastle & A. Marchal,
745 (Academic Publishers), 403-439.

746 Lindblom, B., (1990b). "On the notion of "possible speech sound"," J. Phonetic, **18**, 135-
747 152.

748 Lindblom, B., MacNeilage, P.F. and Studdert-Kennedy, M. (1984). *Self-organizing*
749 *Processes and the Explanation of Languages Universals*, in *Explanations of*
750 *Linguistic Universals*, edited by Butterworth, B., Comrie, B., and Dahl, O., (Mouton),
751 181-203.

752 Locke, J. L. (2000). "Movement patterns in spoken language," *Science*, **288**, 449-451.

753 Locke, J. L. (1983). *Phonological Acquisition and Change*, (Academic Press, New York).

754 MacNeilage, P.F. (1998). "The frame/content theory of evolution of speech production,"
755 *Behav. Brain Sci.* **21**, 499-511.

756 MacNeilage, P.F., Davis, B.L., Kinney, A. and Matyear, C.L. (2000a). "The motor core of
757 speech: A comparison of serial organizations patterns in infants and languages,"
758 *Child. Dev.* **71**, 153-163.

759 MacNeilage, P.F. and Davis, B.L. (2000b). "On the origins of intenal structure of word
760 forms," *Science*, **288**, 527-531.

761 Munhall, K.G. and Jones, J.A. (1998). "Articulatory evidence for syllabic structure,"
762 *Behav. Brain Sci.* **21**, 524-525.

763 Nelson, W. L. (1983). "Physical principles for economies of skilled movements," *Biol.*
764 *Cybern.* **46**, 135-147.

765 Nelson, W.L., Perkell, J.L. and Westbury, J.R. (1984). "Mandibule movements during
766 increasingly rapid articulations of single syllables : Preliminary observations," J.
767 Acoust. Soc. Am. **75**, 945-951.

768 Nittrouer, S. (1991). "Phase relations of jaw and tongue tip movements in the production
769 of VCV utterances," J. Acoust. Soc. Am. **90**, 1806-1815.

770 Ostry, D.J. and Flanagan, J.R. (1989). "Human jaw movement in mastication and
771 speech," Arch. Oral Biol., **34**, 685-693.

772 Oudeyer, P-Y. (2001). "The origins of syllable systems: An operational model," in
773 *Proceedings of the Twenty-third Annual Conference of the Cognitive Science Society*,
774 edited by J. D. Moore and K. Stenning (Lawrence Erlbaum Associates), 744-749.

775 Perkell, J. S., Zandipour, M., Matthies, M. L. and Lane, H. (2002). "Economy of effort in
776 different speaking conditions. I. A preliminary study of intersubject differences and
777 modeling issues," J. Acoust. Soc. Am. **112**, 1627-1641.

778 Perkell, J. S. and Zandipour, M. (2002). "Economy of effort in different speaking
779 conditions. II. kinematic performance spaces for cyclical and speech movements," J.
780 Acoust. Soc. Am. **112**, 1642-1651.

781 Rhardiss, N. and Abry, C. (1994). "La coarticulation mandibulaire comme principe
782 d'organisation syllabique (Et si la mandibule syllabisait ALIS?)," in Proceedings of
783 the XXèmes Journée d'Etude sur la Parole, Trégastel, 421-426.

784 Redford, M.A., Chen, C.C. and Miikkulainen, R. (2001). "Constrained emergence of
785 universals and variation in syllable systems," *Language and Speech*, **44**, 27-56.

786 Redford, M.A. and Diehl, R.L. (1999). "The relative perceptual distinctiveness of initial
787 and final consonants in cvc syllables," J. Acoust. Soc. Am. **103**, 1555-1565.

788 Rousset, I. (2003). "From lexical to syllabic organization: favored and disfavored co-
789 occurrences," In *Proceedings of the 15th ICPHS* (Barcelona, Spain), 715-718.

790 Sato, M., Schwartz, J-L., Cathiard, M.A., Abry, C. and Loevenbruck, H. (2006).
791 "Multistable syllables as enacted percepts: a source of an asymmetric bias in the
792 verbal transformation effect," *Perception & Psychophysics*, **68** (3), 458-474.

793 Selkirk, E. (1984). *On the major class features and syllable theory* (MIT Press).

794 Sonoda, Y. and Nakakido, K. (1986). "Effetct of Speaking Rate on Jaw Movements in
795 Vowel Sequence" J. Acoust. Soc. Jpn (E) **7**, 5-12.

796 Stetson, R. H. (1951). *Motor Phonetics* (North-Holland, Amsterdam), 2nd ed.

797 Stevens, K. N. (1989). "On the Quantal Nature of Speech," J. Phonetics, **17**, 3 - 45.

798 Surprenant, A. M. and Goldstein, L. (1998). "The perception of speech gestures," J.
799 Acoust. Soc. Am., **104** (1), 518-529.

800 Tuller, B. and Kelso, J. A. S. (1991). "The production and perception of syllable
801 structure," J. Speech Hear. Res. **34**, 501-508.

802 Vilain, A., Abry, C., Badin, P. and Brosda, S. (1999). "From idiosyncratic pure frame to
803 variegated babbling: evidence from articulatory modelling," in *Prodeedings of*
804 *ICPhS'99* (San Fransisco, USA), 2497-2500.

805 Zsiga, E. C. (1996). "Acoustic evidence for gestural overlap in consonant sequence," J.
806 Phonetics, **22**, 121-140.

807

807 Tables

808

808

809 *Table I: Analysis of intensity variations between the vowel after the labial and the*
810 *vowel after the coronal consonant for utterances shorter than 300 ms for the 42 analyzed*
811 *trials (see text for the details about data selection). For each disyllable: means (dB) and*
812 *are compared with 0 according to a two-tails t-test (df = 41) with Bonferroni correction*
813 *(* $p \leq .01/6$); standard errors; repartition of the 42 means according to their sign: for all*
814 *the means and for the means that significantly differ from 0 according to a two-tails t-test*
815 *with Bonferroni correction ($p \leq .05/42$).*

		pata	tapa	pasa	sapa	fata	tafa
Means		6.6 *	7.0 *	7.5 *	3.8	8.9 *	8.6 *
Standard errors		1.9	1.9	1.7	1.7	1.7	2.1
Number of positive means	All	30	30	27	21	37	33
	Signif.	19	15	18	13	26	21
Number of negative means	All	12	12	15	21	5	9
	Signif.	4	6	6	9	2	4

816

817

817 *Tableau II: Number of CVCV or CCV utterances correctly produced (Tot.) and number of*
 818 *utterances produced with sufficient jaw motion (see text) and kept for further articulatory*
 819 *analysis (Ar.) for each speaker and each sequence.*

	papa		tata		pata		tapa		pasa		sapa		fata		tafa	
	Tot.	Ar.	Tot.	Ar.	Tot.	Ar.	Tot.	Ar.	Tot.	Ar.	Tot.	Ar.	Tot.	Ar.	Tot.	Ar.
S1	63	63	63	60	61	59	59	54	60	59	54	54	53	51	51	47
S2	51	49	53	52	58	47	52	48	62	43	59	43	65	59	58	49
S3	53	51	56	54	60	59	63	61	59	59	64	64	57	57	65	65
S4	66	63	68	60	76	52	73	55	74	73	62	60	69	64	71	53
S5	94	85	100	76	109	93	102	80	99	73	95	82	98	79	100	74

820

820 *Tableau III: Means (m) of intensity variation between the vowel after the coronal*
 821 *and the vowel after the labial consonant for utterances shorter than 300 ms for the three*
 822 *trials taken together, for each speaker and each disyllable (n is the number of observations*
 823 *for m computation). Thirty t-tests (df = n-1) with Bonferroni correction compare each*
 824 *mean to 0 (° p ≤ .005, * p ≤ .05/30, ** p ≤ .01/30).*

	pata		tapa		pasa		sapa		fata		tafa	
	n	m	n	m	n	m	n	m	n	m	n	m
S1	31	0.8	29	-0.1	29	-1.4*	27	-2.9**	23	-0.2	24	-0.7
S2	25	23.2**	11	9.2	25	11.7*	20	5.5	30	20.3**	20	13.9**
S3	5	-0.8	9	-0.7	0	-	0	-	0	-	10	1.5
S4	43	-0.7	37	-4.7°	37	-0.1	30	-2.4**	28	0.3	29	-1.6°
S5	67	6.9**	59	4.7**	56	1.7°	45	-2.5**	54	7.6**	53	1.1

825

826

827

828

828 *Tableau IV: Means of amplitudes and durations for the jaw cycles that come with*
 829 *the labial and the coronal constrictions in the acceleration phase, for each speaker and*
 830 *each L_aC_o / C_oL_a pair (n is the number of observations for means computations). The*
 831 *significance of the difference between means was tested using t -tests ($df = n-1$) with*
 832 *Bonferroni corrections (** $p \leq .01/15$, * $p \leq .05/15$, NS : Not Significant).*

Speaker	Sequence	n	Amplitudes (mm)			Durations (ms)		
			labial	coronal	p	labial	coronal	p
s1	pata/tapa	47	0,58	0,91	**	163,6	198,6	**
	pasa/sapa	38	0,59	2,06	**	162,7	247,2	**
	fata/tafa	23	0,98	0,77	NS	272,6	222,1	NS
s2	pata/tapa	21	1,83	1,57	NS	231,9	244,0	NS
	pasa/sapa	28	1,07	1,48	**	209,1	285,2	**
	fata/tafa	23	1,00	0,83	NS	263,7	235,7	NS
s3	pata/tapa	59	1,17	2,75	**	178,3	229,7	**
	pasa/sapa	55	1,01	3,46	**	169,9	269,5	**
	fata/tafa	54	2,02	2,36	**	211,2	219,5	NS
s4	pata/tapa	23	1,16	2,67	**	174,7	247,0	**
	pasa/sapa	18	1,20	3,89	**	206,1	279,2	NS
	fata/tafa	33	1,57	1,34	NS	223,3	214,6	NS
s5	pata/tapa	35	0,85	2,53	**	204,7	268,0	**
	pasa/sapa	14	1,08	3,78	**	196,0	370,7	**
	fata/tafa	52	1,43	1,66	NS	267,8	247,0	NS

833

834

834 *Tableau V: Labial-to-coronal duration means (m) normalized by the whole*
835 *utterance duration $\Delta(L_aC_o/L_aL_a)$, expressed in percent, for each speaker and each disyllable*
836 *(n is the number of utterances involved in computing m). Values smaller than 50% indicate*
837 *that the duration from the labial constriction to the following coronal one is smaller than*
838 *the duration from the coronal constriction to the following labial one, which signals an*
839 *evolution towards the L_aC_o attractor. On the contrary, values higher than 50% signal an*
840 *evolution towards the C_oL_a attractor. Thirty two-tails t-tests ($df = n-1$) with Bonferroni*
841 *correction compare each mean to 50 (* $p \leq .05/30$, ** $p \leq .01/30$).*

842

	pata		tapa		pasa		sapa		fata		tafa	
	n	m	n	m	n	m	n	m	n	m	n	m
S1	20	49,3	5	53,0	27	50,8	26	51,0	31	53,7**	32	52,0
S2	25	36,3**	21	40,3*	20	41,1	22	43,9	40	30,5**	36	36,1**
S3	19	44,3*	11	38,6**	22	41,5**	17	42,5**	10	44,8	23	43,7**
S4	56	45,4*	48	49,5	60	46,3**	44	52,0	38	42,8**	30	51,4
S5	67	42,1**	58	40,8**	67	51,8	71	53,5*	35	37,9**	41	44,9*

843

844

845

846

846 List of figures

847 *Figure 1: Labeling of intensity curves and measurements: intensity curves (bottom) of the acoustic signal (top)*
 848 *against time for three samples of a /tapa/ trial. The repetition progressively evolves from /tapa/ (left) towards /tápa/*
 849 *(middle) until /ptá/ (right). L_a and C_o are the labial and coronal consonants; V_{L_a} and V_{C_o} are the vowels respectively*
 850 *following L_a and C_o . This labeling allows computing the duration of an utterance and the intensity variation of*
 851 *intensity between V_{L_a} and V_{C_o} (ΔI)..... 49*

852 *Figure 2: Intensity variation between the vowel after the labial constriction and the vowel after the coronal*
 853 *constriction (ΔI) plotted against duration for /pasa/ (left) and /sapa/ (right) CVCV and CCV correctly produced*
 854 *utterances for all subjects (top). Results for /pasa/ and /sapa/ are representative of the two other L_aC_o (/pata/, /fata/)*
 855 *and C_oL_a (/tapa/, /tafa/) pairs. Bottom: Standard deviations of ΔI against utterances durations (100 ms meantime)*
 856 *for the 6 CVCV sequences. Dispersion increases when duration becomes shorter than 300 ms..... 50*

857 *Figure 3: Attractor stability and speakers' profiles. **On the left:** ΔI mean (variation of intensity between the vowel*
 858 *after the labial consonant and the vowel after the coronal consonant) of the second trial against ΔI mean of the first*
 859 *trial for each CVCV and each speaker when the two means significantly differ from zero (two-tails Student test with*
 860 *Bonferroni correction $p \leq .05/42$). Points inside the top left and the bottom right squares correspond to cases with*
 861 *attractor shift from one trial to the other. Points inside the top right and bottom left squares respectively represent*
 862 *L_aC_o and C_oL_a stability for the two trials. **On the right:** ΔI means for the two selected trials for each CVCV of four*
 863 *typical speakers with S20 and S6 as bi-stable profiles; S12 as L_aC_o stable and S23 as C_oL_a stable. The convex hulls of*
 864 *S23, S20 and S12 are displayed on the figure on the left (* : mean significantly different from zero, $p \leq .05/42$). 51*

865 *Figure 4: Acoustic signal, jaw, Tongue Tip (TT) and Lower Lip (LL) displacements against time for samples of a*
 866 */tapa/ trial (S3). From left to right: sample of the acceleration phase; sample of the period with one jaw cycle for two*
 867 *syllables (splitting) and sample of the deceleration phase. The arrows on trajectory curves represent the labeled*
 868 *minima and maxima for the jaw and the maxima for the constrictors. Notice the secondary maxima for TT and LL*
 869 *due to passive motion induced by the jaw gesture that comes with the other constriction (these secondary maxima*
 870 *are not incorporated in further analyses in Experiment 2). 52*

871 *Figure 5: Articulatory measurements. A: duration and amplitude of the jaw cycle are respectively the sum of its up*
872 *and down stroke durations and the mean of its up and down stroke amplitudes. B: position of the Tongue Tip (P(TT))*
873 *and of the Lower Lip (P(LL)); constriction events were displayed on the jaw cycle normalized between -50% and*
874 *50% with -50% at the beginning of the up stroke, 0% at the jaw maxima and 50% at the end of the down stroke. 53*

875 *Figure 6: Intensity variation between the vowel after the labial constriction and the vowel after the coronal*
876 *constriction (ΔI) against utterance duration for all subjects' duplicated (left) and $L_aC_o - C_oL_a$ (right) productions. ΔI*
877 *does not vary much with duration for duplicated logatoms whereas duration around 300 ms (vertical dotted line)*
878 *appears as a boundary for the range of possible ΔI values for L_aC_o and C_oL_a logatoms (see text). 54*

879 *Figure 7: First row: Jaw cycle durations against utterance durations for all speakers' productions and separately*
880 *for duplicated (left) and L_aC_o / C_oL_a sequences (right). Distribution splitting observed for the L_aC_o and C_oL_a groups*
881 *shows that these disyllables could be produced on a single jaw cycle whereas duplicated disyllables are always*
882 *realized on two jaw cycles. Second and third rows: distribution of utterance and jaw cycle durations for all speakers'*
883 *productions. The utterance durations tend to be shorter for the L_aC_o / C_oL_a group as compared to the duplicated one*
884 *whereas the contrary tends to appear for the jaw cycle durations. 55*

885 *Figure 8: Jaw strokes amplitude (up and down) against duration during the acceleration phase (when the disyllables*
886 *are realized with two jaw cycles) for each speaker and separately for duplicated (left column) and L_aC_o / C_oL_a (right*
887 *column) disyllables. The parabolic curves are the theoretical minimum-time boundaries for mandible movements for*
888 *three control acceleration limits (from left to right: 1.5 g, 0.5 g and 0.25 g, (see Nelson 1983)). Regression lines and*
889 *distribution means (squares and circles) are displayed separately for jaw strokes that come with labials and with*
890 *coronals (see text for details). 56*

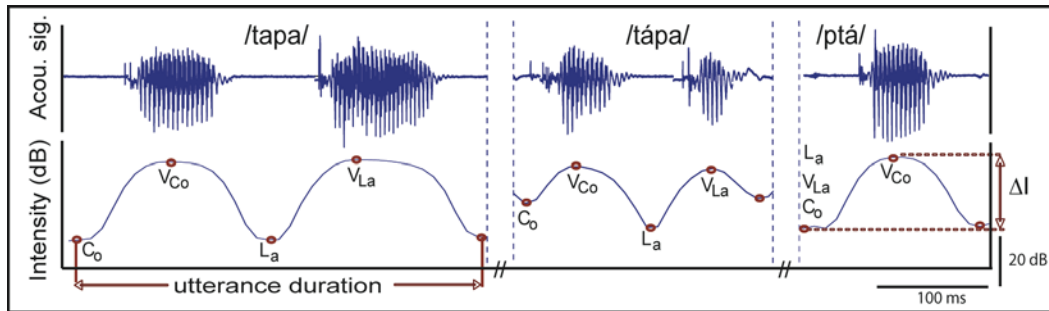
891 *Figure 9: Distribution of speakers' productions according to the position of the Lower-Lip (left) and the Tongue-Tip*
892 *(right) constrictions relative to the jaw cycle. The three phases (acceleration, splitting and deceleration) are plotted*
893 *separately. For the acceleration and the deceleration phases, histograms include all speakers taken together and*
894 *both duplicated and L_aC_o / C_oL_a sequences. For the splitting phase (when utterances are realized on a single jaw*
895 *cycle), distributions are given for each L_aC_o and C_oL_a sequence and for each speaker. 57*

896 *Figure 10: Coordination between the jaw and the constrictors from two jaw cycles for a disyllable to a single one.*
897 *Two patterns (A and B) of lower lip (LL) and tongue tip (TT) constriction positions relative to the jaw cycle are*

898 *observed when the two plosives are realized on a single jaw cycle (splitting, right). They tend to correspond to*
899 *specific jaw motion profiles during the acceleration phase (left, see text for detail). 58*
900

900 Figures

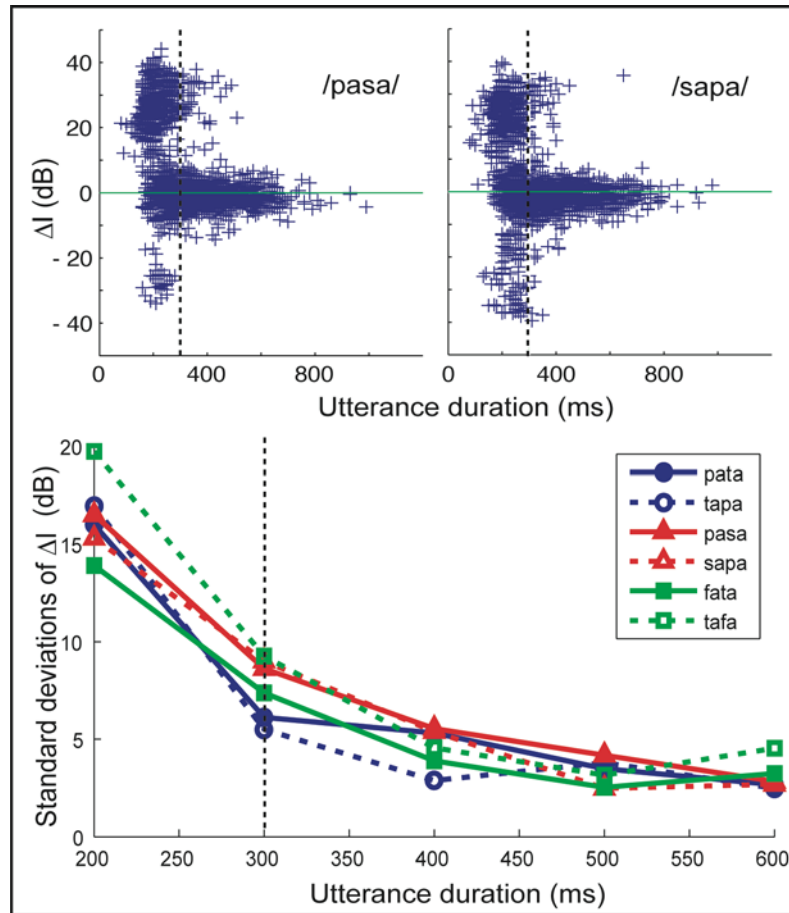
901



901

902 *Figure 1: Labeling of intensity curves and measurements: intensity curves (bottom) of the*
 903 *acoustic signal (top) against time for three samples of a /tapa/ trial. The repetition*
 904 *progressively evolves from /tapa/ (left) towards /tápa/ (middle) until /ptá/ (right). L_a and C_o*
 905 *are the labial and coronal consonants; V_{La} and V_{Co} are the vowels respectively following L_a*
 906 *and C_o. This labeling allows computing the duration of an utterance and the intensity*
 907 *variation of intensity between V_{La} and V_{Co} (ΔI).*

908



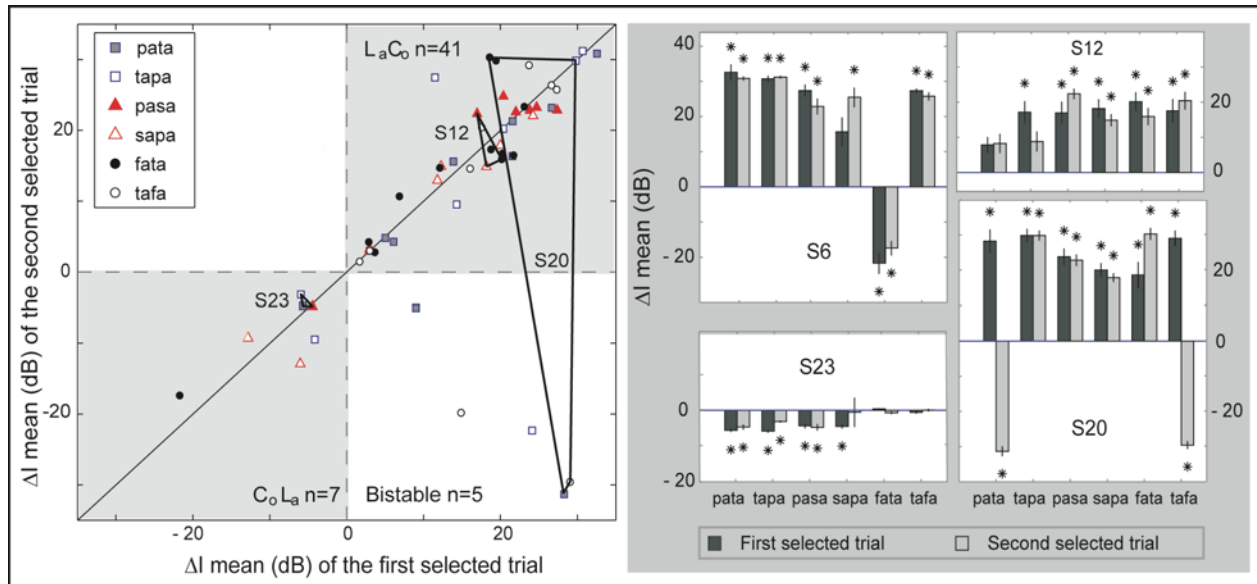
908

909 *Figure 2: Intensity variation between the vowel after the labial constriction and the vowel*
 910 *after the coronal constriction (ΔI) plotted against duration for /pasa/ (left) and /sapa/ (right)*
 911 *CVCV and CCV correctly produced utterances for all subjects (top). Results for /pasa/ and*
 912 */sapa/ are representative of the two other L_aC_o (/pata/, /fata/) and C_oL_a (/tapa/, /tafa/) pairs.*
 913 *Bottom: Standard deviations of ΔI against utterances durations (100 ms meantime) for the 6*
 914 *CVCV sequences. Dispersion increases when duration becomes shorter than 300 ms.*

915

916

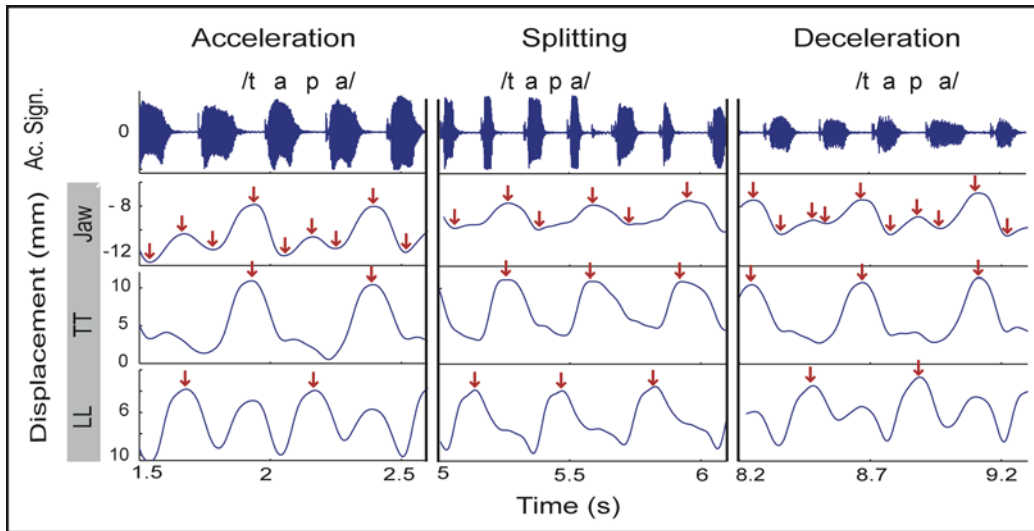
917



917

918 *Figure 3: Attractor stability and speakers' profiles. **On the left:** ΔI mean (variation of*
 919 *intensity between the vowel after the labial consonant and the vowel after the coronal*
 920 *consonant) of the second trial against ΔI mean of the first trial for each CVCV and each*
 921 *speaker when the two means significantly differ from zero (two-tails Student test with*
 922 *Bonferroni correction $p \leq .05/42$). Points inside the top left and the bottom right squares*
 923 *correspond to cases with attractor shift from one trial to the other. Points inside the top*
 924 *right and bottom left squares respectively represent L_aC_o and C_oL_a stability for the two*
 925 *trials. **On the right:** ΔI means for the two selected trials for each CVCV of four typical*
 926 *speakers with S20 and S6 as bi-stable profiles; S12 as L_aC_o stable and S23 as C_oL_a stable.*
 927 *The convex hulls of S23, S20 and S12 are displayed on the figure on the left (* : mean*
 928 *significantly different from zero, $p \leq .05/42$).*

929

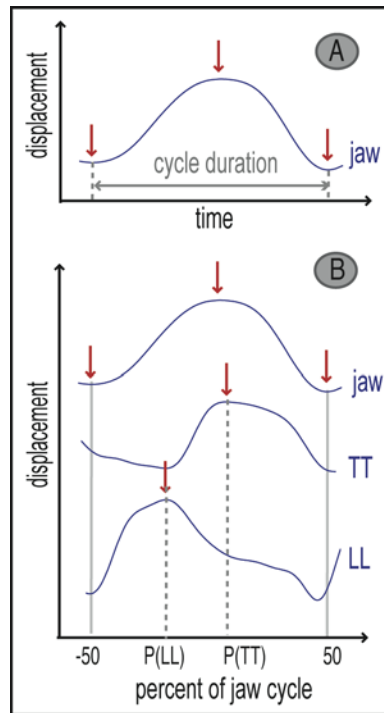


929

930 *Figure 4: Acoustic signal, jaw, Tongue Tip (TT) and Lower Lip (LL) displacements against*
 931 *time for samples of a /tapa/ trial (S3). From left to right: sample of the acceleration phase;*
 932 *sample of the period with one jaw cycle for two syllables (splitting) and sample of the*
 933 *deceleration phase. The arrows on trajectory curves represent the labeled minima and*
 934 *maxima for the jaw and the maxima for the constrictors. Notice the secondary maxima for*
 935 *TT and LL due to passive motion induced by the jaw gesture that comes with the other*
 936 *constriction (these secondary maxima are not incorporated in further analyses in*
 937 *Experiment 2).*

938

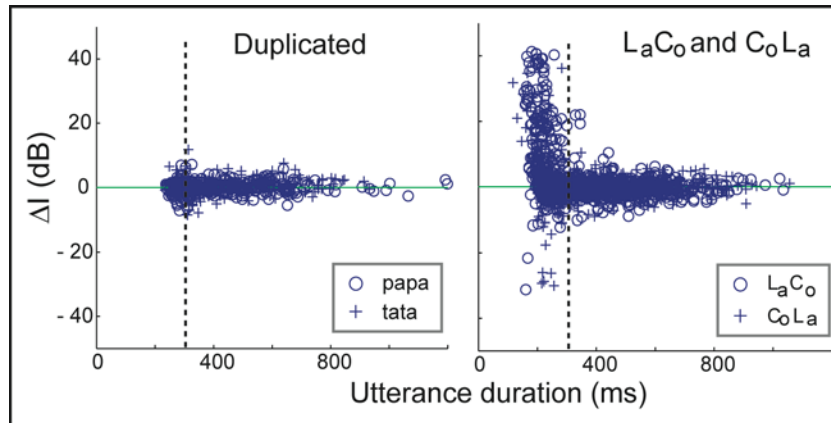
938



939

940 *Figure 5: Articulatory measurements. A: duration and amplitude of the jaw cycle are*
941 *respectively the sum of its up and down stroke durations and the mean of its up and down*
942 *stroke amplitudes. B: position of the Tongue Tip (P(TT)) and of the Lower Lip (P(LL));*
943 *constriction events were displayed on the jaw cycle normalized between -50% and 50%*
944 *with -50% at the beginning of the up stroke, 0% at the jaw maxima and 50% at the end of*
945 *the down stroke.*

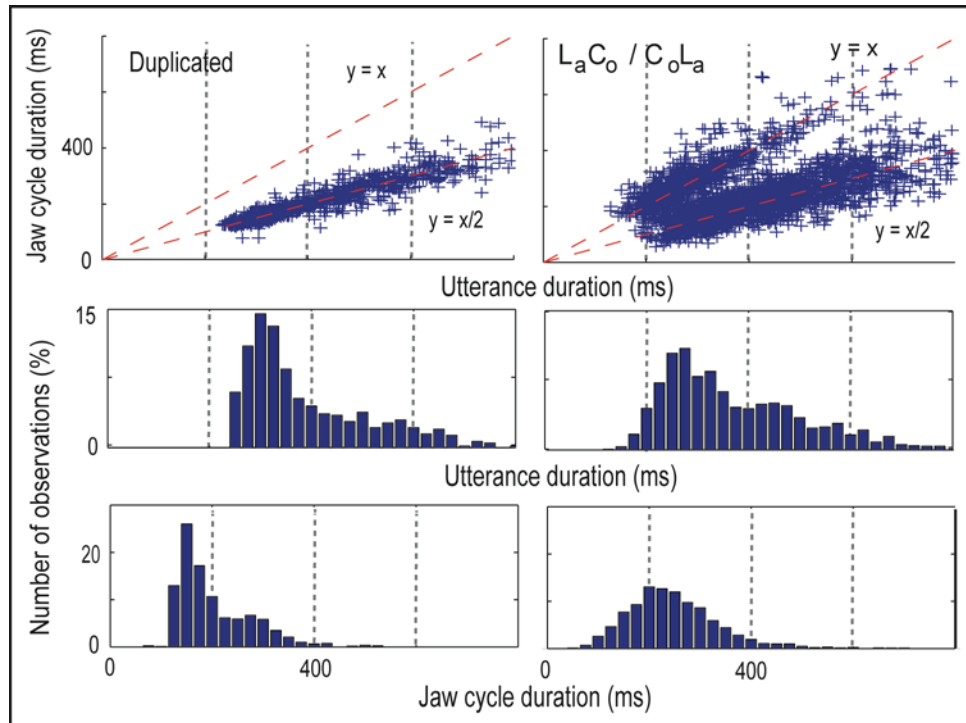
946



946

947 *Figure 6: Intensity variation between the vowel after the labial constriction and the vowel*
 948 *after the coronal constriction (ΔI) against utterance duration for all subjects' duplicated*
 949 *(left) and L_aC_o - C_oL_a (right) productions. ΔI does not vary much with duration for*
 950 *duplicated logatons whereas duration around 300 ms (vertical dotted line) appears as a*
 951 *boundary for the range of possible ΔI values for L_aC_o and C_oL_a logatons (see text).*

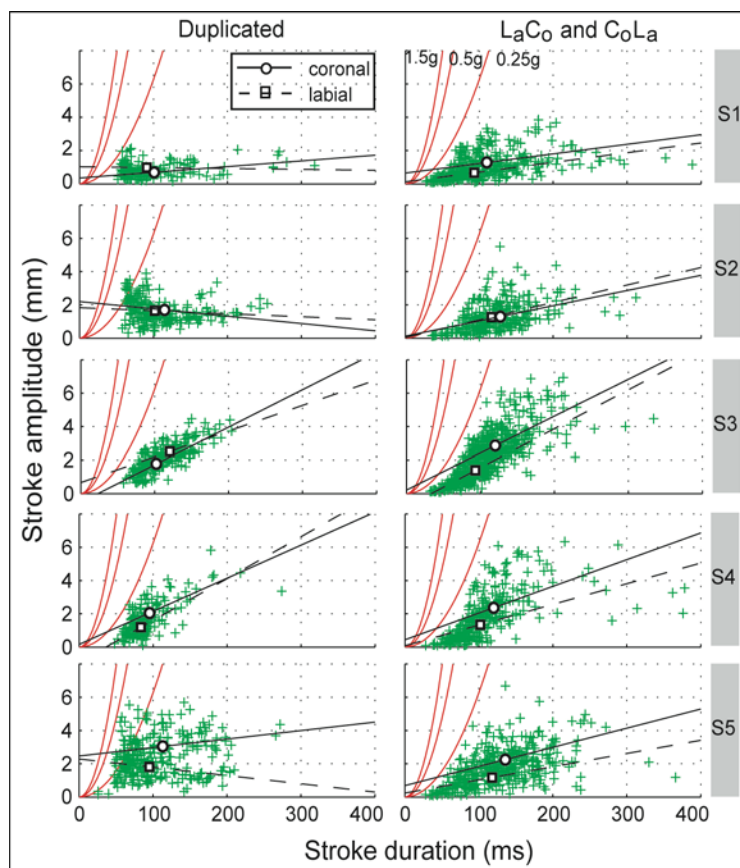
952



952

953 *Figure 7: First row: Jaw cycle durations against utterance durations for all speakers'*
 954 *productions and separately for duplicated (left) and L_aC_o / C_oL_a sequences (right).*
 955 *Distribution splitting observed for the L_aC_o and C_oL_a groups shows that these disyllables*
 956 *could be produced on a single jaw cycle whereas duplicated disyllables are always realized*
 957 *on two jaw cycles. Second and third rows: distribution of utterance and jaw cycle durations*
 958 *for all speakers' productions. The utterance durations tend to be shorter for the L_aC_o / C_oL_a*
 959 *group as compared to the duplicated one whereas the contrary tends to appear for the jaw*
 960 *cycle durations.*

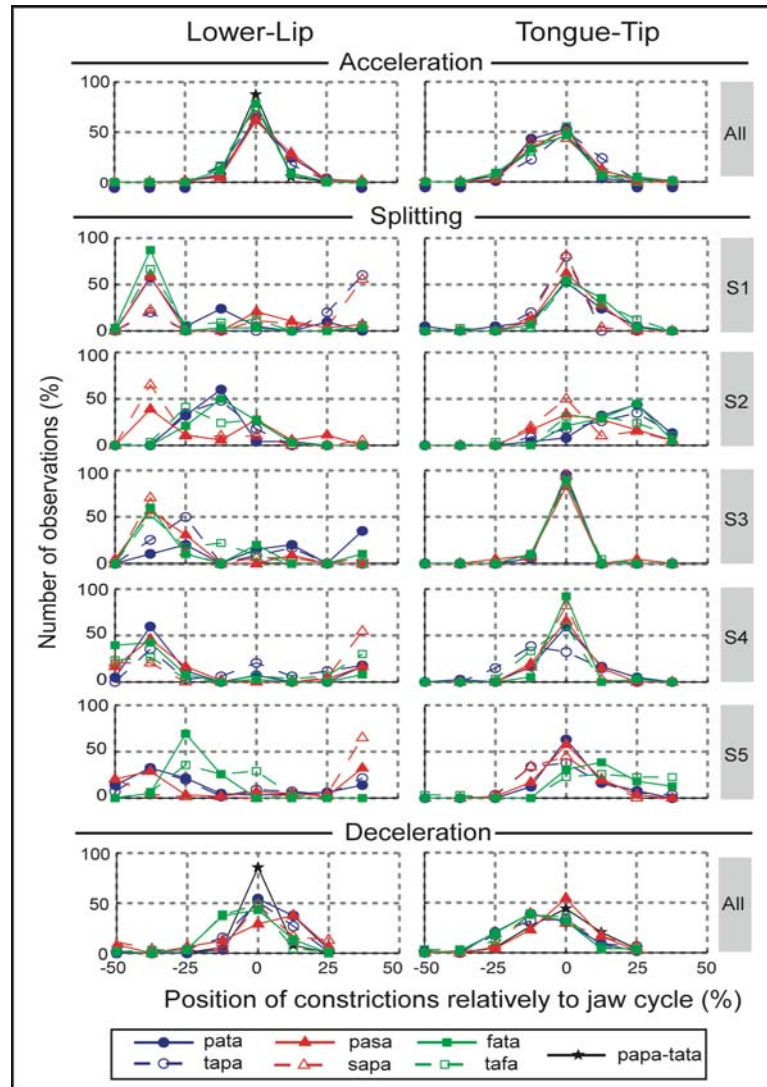
961



961

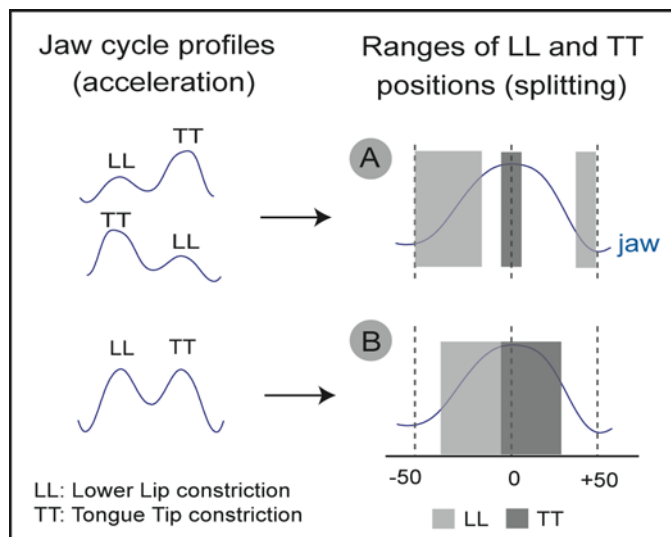
962 *Figure 8: Jaw strokes amplitude (up and down) against duration during the acceleration*
 963 *phase (when the disyllables are realized with two jaw cycles) for each speaker and*
 964 *separately for duplicated (left column) and L_aC_o / C_oL_a (right column) disyllables. The*
 965 *parabolic curves are the theoretical minimum-time boundaries for mandible movements for*
 966 *three control acceleration limits (from left to right: 1.5 g, 0.5 g and 0.25 g, (see Nelson*
 967 *1983)). Regression lines and distribution means (squares and circles) are displayed*
 968 *separately for jaw strokes that come with labials and with coronals (see text for details).*

969



969

970 *Figure 9: Distribution of speakers' productions according to the position of the Lower-Lip*
 971 *(left) and the Tongue-Tip (right) constrictions relative to the jaw cycle. The three phases*
 972 *(acceleration, splitting and deceleration) are plotted separately. For the acceleration and*
 973 *the deceleration phases, histograms include all speakers taken together and both*
 974 *duplicated and L_aC_o / C_oL_a sequences. For the splitting phase (when utterances are realized*
 975 *on a single jaw cycle), distributions are given for each L_aC_o and C_oL_a sequence and for*
 976 *each speaker.*



977

978 *Figure 10: Coordination between the jaw and the constrictors from two jaw cycles for a*
 979 *disyllable to a single one. Two patterns (A and B) of lower lip (LL) and tongue tip (TT)*
 980 *constriction positions relative to the jaw cycle are observed when the two plosives are*
 981 *realized on a single jaw cycle (splitting, right). They tend to correspond to specific jaw*
 982 *motion profiles during the acceleration phase (left, see text for detail).*

983

984

