Articulatory and acoustic correlates of contrastive focus in French children and adults

Lucie Ménard, Hélène Loevenbruck, Christophe Savariaux

Articulatory and acoustic correlates of contrastive focus in French children and adults

Lucie Ménard[a], *Hélène Lœvenbruck*[b], *and Christophe Savariaux*[b]

[a]Université du Québec à Montréal, Canada

[b]Institut de la Communication Parlée, France

menard.lucie@uqam.ca, loeven@icp.inpg.fr, savario@icp.inpg.fr

This chapter deals with the articulatory and acoustic correlates of contrastive focus in French, in children and adults. Ten speakers (three 4-year-old, three 8-year-old children, and four adult males and females) were recorded while producing repetitions of the [baba] sequence in two prosodic contexts: in neutral condition and under contrastive focus. The opening (from the consonant into the vowel) lip gestures were video-recorded and monitored using the ICP tracking system. Lip area and its first derivative were described under both prosodic conditions, in the first and second [ba] syllables. Analyses of formant frequencies and RMS amplitude were also carried out on the acoustic signal. Results show that the effect of contrastive focus on articulatory measurements is smaller for children compared to adults. Results are discussed with respect to motor control development during childhood.

INTRODUCTION

The development of speech entails the control of various components of the vocal tract. The sequential mastery of the larynx, jaw, and tongue, for instance, shape the vocal repertoire of the young speaker (Vihman, 1996). Learning to speak also means embedding a sequence of segments within a hierarchically organized prosodic structure. It has been shown in various languages that adult speakers signal prosodic structure by acoustic and articulatory cues both realized by the glottal and supraglottal articulators (e.g. Beckman *et al.*, 1992). As detailed below, in English and French, for example, it is reported that syllables produced with contrastive focus are uttered with increased values of fundamental frequency (F0), duration, and/or intensity, compared to their unfocused counterparts (e.g. Ladd, 1996; Selkirk, 1984; Touati, 1987), and are also associated with changes in the kinematics of the supralaryngeal articulators (e.g. De Jong, 1995; Dohen *et al*., 2004b; Erickson, 2002; Lœvenbruck, 1999). In typically-developing children, glottal and supraglottal mechanisms are controlled at different stages during development (Vihman, 1996). Hence it is possible that different articulatory strategies are involved in prosodic marking for adults and children. This paper aims at examining production differences between 4-year-old, 8-year-old, and adult French speakers in marking focus-induced prominence at the articulatory and acoustic levels.

CONTRASTIVE FOCUS IN ADULTS

Contrastive focus is one type of narrow focus defined as involving the selection by the speaker of a given constituent of the message to be underlined as opposed to another constituent in a paradigmatic comparison (Bartels & Kingston, 1994; Dahan & Bernard, 1996; Di Cristo, 2000; Ladd, 1996; Pierrehumbert & Hirshberg, 1990; Selkirk, 1984; Touati, 1987). In French, contrastive focus can either be signalled syntactically (the prominent word or phrase X is put forward through a syntactic focus construction, such as a syntactic extraction: “c’est X qui”, “it’s X who”) or prosodically (the prominent word or phrase is put forward through the manipulation of prosodic features such as fundamental frequency (F0), duration and/or intensity), or both (e.g.

Di Cristo, 2000). The present paper deals with contrastive focus conveyed by both syntax and prosody, which will be referred to simply as “contrastive focus”. Studies carried out in English have shown that this prosodic prominence is associated with larger, faster, and longer lip opening gestures for /a/ and /i/ (Cho, 2002; Cho, in press) and lower jaw positions for /a/ (Erickson, 1998). Lœvenbruck (1999) also reports that tongue position is lower for [la] syllables produced under contrastive focus, compared to the neutral condition. These articulatory correlates are often related to spectral changes, with low vowels being more peripheral in the acoustic space (Cho, in press). It is important to note that between-speaker variability is reported with respect to the articulatory and acoustic effects of prominence.

This study is concerned with contrastive focus in French. It is important to note that unlike English, French is usually described as a fixed-accent language in which an obligatory primary accent is assigned to the final full syllable of a prosodic phrase. This demarcative accent, referred to as H* in Jun & Fougeron (2000, 2002), marks the right edge of a prosodic phrase (see also e.g. Di Cristo, 2000; Jun and Fougeron, 2000, 2002; Lacheret-Dujour & Beaugendre, 1999; Pasdeloup, 1990). When carrying a primary accent, in French, syllables are longer and higher in intensity than the preceding syllables. In addition, if the phrase is not utterance final, the primary-accented syllable bears a rise in F0. In addition to the phrase-final primary accent, French features an optional phrase-initial secondary accent, referred to as Hi in Jun & Fougeron (2000, 2002). The location of the secondary accent is variable: it generally occurs on one of the first syllables of the phrase (see Welby, 2003 for data and hypotheses on the alignment of the initial rise). Contrary to the primary accent, the secondary accent is not always accompanied by syllabic lengthening or increased intensity and its associated F0 peak is generally lower than that of the primary accent.

The intonational correlates of contrastive focus in French have been widely studied (e.g. Clech-Darbon *et al.*, 1999; Delais-Roussarie *et al*., 2002, Di Cristo, 1998; Dohen *et al*., 2004; Jun & Fougeron, 2000; Rossi, 1999; Touati, 1987). They include a large and sharp rise in F0 and/or intensity on the focused constituent, an increased duration of the focused syllables and a global F0 and intensity compression in the post-focus sequence, with either a low plateau or a steady fall until the end of the utterance. The F0 peak, or Hf, most often replaces Hi; that is, Hf is usually located on one of the initial syllables. But Hf can also sometimes replace H* or even both Hi and H* (the rise in F0 can be born by all the syllables and culminates on the last syllable). The post-focal sequence is usually deaccented, i.e. it is realized as a low-plateau in which all tones are deleted but phrase boundary lengthening is still present (no dephrasing).

Although the terminology used may differ, a number of articulatory studies deal with what we consider as contrastive focus, in several languages. In order to explain the effects of contrastive focus on the kinematic patterns of the supraglottal articulators, two models are proposed. In a first model, the mechanism involved in the variation of the supraglottal articulation associated with the prominence would be a local hyperarticulation of the focused syllable, evidenced by reduced coarticulation (de Jong, 1995). A similar hyperarticulation account is also proposed by Lœvenbruck (1999) for the tongue movements, in a study of the strategies used by two speakers to achieve contrastive focus in French. Note however that a recent study by Erickson, Iskarous, and Whalen (2005) has shown that hyperarticulation and contrastive focus differ in terms of the relationship between the jaw and the tongue. In a task involving hyperarticulation of vowels in English, the tongue and jaw were lower and positively correlated. On the contrary, in the contrastive focus task, no linear relationship was found between the tongue and the jaw, the tongue moving independently from the jaw. According to this result, hyperarticulation and contrastive focus do not involve the same articulatory reorganization. In a second account of the articulatory correlates of contrastive focus, the speaker would lower the jaw and the tongue dorsum in order to increase the area of the oral passage, and thus, to increase the loudness of the vowel. This model is referred to as the "sonority expansion" model (Beckman *et al.*, 1992). For low vowels, both hypotheses lead to the same predicted articulatory correlates: lowering of the jaw and the tongue, and larger lip area**.**

THE DEVELOPMENT OF PROSODIC CORRELATES

From a developmental point of view, since the laryngeal and supralaryngeal mechanisms involved in the focused vs. unfocused contrast are mastered at different growth stages (Vihman, 1996), it is possible that children recruit different articulatory strategies compared to adults when signalling contrastive focus. For instance, it is likely that the supralaryngeal articulators are not as finely tuned in young children (in speed and/or magnitude of displacement) as they are in adults. Yong children would then not be able to exploit the full scale articulatory dimension, like adults seem to do, as suggested by the hyperarticulation hypothesis discussed above. If this were the case, adult-like articulatory and acoustic correlates of contrastive focus would gradually emerge with age.

The ability to signal contrastive focus using syntax and prosody has been found in children from 3 years old (Hornby and Hass, 1970). However, correlates of prosodic prominence have mainly been studied, in English, for lexical stress. For instance, Connaghan *et al.* (2001) report that children of 2 years of age and 3.6 years of age use pitch and intensity values comparable to adults in stressed syllables**.** Allen and Hawkins (1980) report that the magnitude of the contrast along each of these acoustic dimensions is reduced for 2 to 4-year-old children, compared to adult speakers. For a review of supporting data, see Kehoe *et al.* (1995). The ability to realize adult-like pitch movements is also observed early in French (Konopczinsky, 1986). In French also, it has been shown that adult-like syllable lengthening is realized by young children (from 2 years of age) on accented vowels (Konopczinsky, 1986).

By contrast, adult-like control of the supraglottal articulators is not achieved until the end of the first decade of life (Green *et al.*, 2002; Vihman, 1996). Besides the overall greater within-speaker variability across repetitions revealed by these studies, the articulatory gestures are realized in children with a reduced velocity, even though they do not differ in magnitude from the adults (Smith and Goffman, 1998). Considering the smaller size of the articulators in children, articulatory movements are expected to have smaller amplitudes. As concerns duration, children have been reported to produce longer segments (Kent and Forner, 1980). Very few studies have attempted to find articulatory correlates of linguistic prominence in children, with the exception of Connaghan *et al.* (2001), in English, who reported that in the production of trochaic and iambic stressed patterns of [baba], lip distance was not correlated with intensity for the children (from 3.6 year old to 6.5 year old), whereas these two parameters were correlated in adults. However, this study was concerned with lexically stressed syllables, and not focused syllables. Goffman (1999) studied the kinematic differentiation of articulatory movements across iambic and trochaic stress patterns in 3-year-old English speaking children. Results showed that children produce adult-like iambs where the first syllable (unstressed) is smaller in amplitude and duration compared to the second syllable (stressed). However, trochees are produced in children with undifferentiated kinematic patterns over both syllables, even though the first one is supposed to be stressed. This pattern contrasts with the adult pattern, for which the first syllable is larger in amplitude and duration compared to the second syllable.

These globally large movements, reduced velocities, and increased durations, for children compared to adults, can be seen as characteristics of slowly and carefully articulated speech. In line with this interpretation, a few studies report an inability in children to reduce vowels in unaccented syllables, therefore leading to a constantly carefully articulated speech style. Allen and Hawkins (1980) studied English-learning children from 1 year to 3 years of age and describe the lexically unstressed vowels as more peripheral than those produced by adults. Thus, Allen and Hawkins (1980) propose that learning to put lexical stress for children means learning to reduce unstressed syllables. In this paper, we test the hypothesis that French-speaking children will not differentiate articulatory and acoustic patterns across accented and unaccented syllables as much as adult speakers will do, under contrastive focus compared to the neutral condition. As a result, the articulatory contrast in lip area as well as the acoustic contrast, in the F1 vs. F2 space, will be reduced for the child. Since we relate the kinematic differentiation between the focus and unfocused conditions to motor control, this study will attempt to test whether supralaryngeal motor control is fully acquired or not by the age of 8.

METHOD

The present study was designed to determine the acoustic and articulatory correlates of contrastive focus in standard metropolitan French, in children and adults. Acoustic correlates include formant values and RMS amplitude. In the articulatory domain, lip area data were obtained.

Speakers and corpus

Six children (three 4-year-olds and three 8-year-olds) and four adult speakers were audio-visually recorded. All subjects were native speakers of French. The six children were evaluated as normal for language development by a vocabulary test (EVIP-P Body). Furthermore, they did not show any hearing or articulatory disability, as evaluated by a brief conversation with the child**.** The target word was [baba], produced in two prosodic conditions: in a neutral condition and under contrastive focus. The target word was embedded in two carrier sentences: “J’ai vu Baba qui mange un gâteau” (*I saw Baba eating a cookie*), in which [baba] is in the neutral condition, and “Non, c’est <u>Baba</u> qui mange un gâteau” (*No, it’s <u>Baba</u> who is eating a cookie*), in which [baba] is focused. These two utterances were part of a story told by the experimenter. In both conditions, the speaker was asked to say what one of the characters of the story had said. In the focus condition, the experimenter asked a question to the speaker, while replacing the target word [baba] by another word. The speaker was asked to correct the wrong utterance by using a syntactic extraction (“No it’s … ”) and producing the word [baba] with contrastive focus[1]. Between five and ten repetitions of each sentence were obtained, for each subject. It is reported

[1] Due to the design of the corpus (which was highly constrained by the fact that we were recording young speakers), contextual effects can be noted. In a pilot study, we had tried to use a similar context for the two conditions, with the sentence “c’est Baba qui mange” in the neutral condition. But this construction generally leads to an unwanted focused intonation on /baba/. We therefore chose to use two different carrier sentences for the two prosodic conditions. As a consequence, the first /a/ in the neutral condition was preceded by a /y/ whereas the first /a/ in the focused condition was preceded by a /e/ or /ɛ/. /y/ and /e, ɛ / differ in the amount of lip protrusion, /y/ being more protruded and thus having a lesser lip area than /e, ɛ/. If there were no inserted consonant between /y/ and the following /a/, persevatory co-articulation effects could certainly influence the amount of protrusion in the /a/, and thus probably decrease lip area for /a/. However /y/ is followed by a closed consonant /b/ which tends to reset the amount of lip protrusion. Although we cannot certify that /y/ has absolutely no influence on the following /a/, with the inserted /b/, we believe that this influence is much reduced compared to the possible influence of prosodic condition.

that deixis, conveyed here by contrastive prosodic focus, is well mastered by preschool children (Hornby and Hass, 1970). In the neutral condition, a high pitch accent (the primary accent) was expected to occur on the second syllable of the target word [baba]. In the contrastive focus condition, a prominence was predicted most likely on the first syllable, but might also occur on the second, or on both syllables of the word [baba], as explained in the previous section. Furthermore, as mentioned previously, it was expected that the second syllable should bear a final lenghtening (compared to the initial syllable) in both conditions. Previous studies have shown that when focus falls on an initial syllable, although the post-focal sequence is deaccented, it is not dephrased, i.e. it keeps durational properties (Jun and Fougeron, 2002). The pitch accented syllable in the neutral condition and the prominent syllable(s) in the focused condition will be simply referred to as "accented" in the following, to remain theory-neutral about their actual prosodic status (pitch accent, insistance accent, focus accent, stress, etc.).

Experimental procedure

Audio and visual data were recorded using the Lip-track system developed at ICP. The images were recorded via one fixed camera and stored on a betacam-ST tape. The speakers wore glasses with two calibration pellets, and sat on a chair. Their head was maintained by a helmet, fixed to the chair. Their lips were painted in blue. The images were digitized at 50 Hz, which corresponds to the standard sampling rate of PAL video format (one video image consists of two interleaved frames). The acoustic signal was also digitized at a frequency of 22050 Hz. Using a program designed at ICP (Audouy, 2000), the external and internal lip contours were automatically detected from the video signal, for the total duration of the target word [baba], after using a chroma-key on blue pixels and filtering the data. Lip area is defined as the area (in mm^2) of the surface inside the internal lip contour. The number of pixels inside the internal contour was thus converted into a surface in cm^2 to obtain lip area (S). The zone shaded by diagonal lines on Figure XXX.1 represents the lip area parameter on a typical video image. Each image was examined by the experimenter, in order to detect computation errors. For some children, the lip contours were hand-corrected, when part of the blue make-up had inadvertently been removed from the lips.

Perceptual analysis of accent patterns

All target words were transcribed by 5 trained phoneticians (four graduate students and one professional phonetician). Following previous studies, each listener had to identify the “accent” (or rather prominence) pattern of the [baba] sequence among three patterns: (i) the first syllable is accented, (ii) the second syllable is accented or (iii) both syllables are accented. For the current study, only sequences perceived by at least 4 judges as bearing an accent on the second syllable, in the neutral condition, and those bearing an accent on the first syllable, in the focus condition, were analysed. This last pattern is the most frequently observed for contrastive focus in French, as explained previously. Table XXX.1 summarizes the number of sentences produced by each speaker in each condition. It can be seen that the neutral condition is always realized with a final accent. In the focused condition, only three speakers produced some occurrences of the target word with an accent on the second syllable. The percentage of tokens produced with an accent on the first syllable is always higher than the reverse pattern (67% for AR, and 83% for SY and NE). Thus, the utterances with a rare prominence pattern (baBA) constituted a very small subset, and since no statistical analysis was possible on this subset, they were not included in the analysis. In all cases, syllables perceived as accented (second syllable in neutral condition and first syllable in focus condition) were produced with a F0 peak. In the focus condition, the typical deaccented F0 pattern on the post-focus sequence was observed. Despite the variation in prominence patterns, all sentences produced in this prosodic condition were perceived as focused by all listeners.

Analysis of articulatory and acoustic data

For each target sequence [baba], the recorded data were analysed as follows. Three synchronous signals were available: the acoustic signal, the evolution of lip area (S), the first derivative of S (referred to as velocity). Figure XXX.2 shows labelled articulatory events, based on minimal values, maximal values and zero-crossing values of S and velocity. For the present study, the following parameters were studied, based on the various events shown in Figure XXX.2:

- the maximal lip area reached for [a] (value of S at the time of v3 and v8)
- the peak velocity of the opening gesture (value of v2 and v7)
- the duration of the opening gesture (the duration between v3 and v1, and v8 and v6)

In addition, formant frequencies at the time of maximal lip area of the vowel were extracted, using LPC algorithms integrated in the Praat software (http://www.fon.hum.uva.nl/praat/). The number of poles varied from 10 to 14. We used a 14-ms Hamming window, with a pre-emphasis factor of 0.98 (pre-emphasis from 50 Hz for a sampling frequency of 22050 Hz). For all vowels, the results were compared to formant values based on a visual inspection of the spectrogram. In case of important discrepancies between the two methods, parameters for the detection algorithm were readjusted and the analysis was performed again. RMS amplitude was also extracted at the time of maximal lip area.

Statistical analyses

First, for each articulatory and acoustic parameter, repeated measures ANOVAs were carried out, with age as the between-subject factor. The prosodic condition (neutral or focus) and the syllable position (first or second syllable) were the within-subject factors. Theoretically, this design requires that each speaker is represented by one cell in each experimental condition. In line with Cho (in press), in this procedure, a factor has a significant effect if all speakers contribute to this effect. It is thus not possible to determine whether a non significant effect is simply due to between-speaker variability, which is often observed in articulatory prosody. Based on Max and Onghena (1999), we included the different trials in each condition as a within-subject factor. This resulted in three-way repeated measures ANOVAs, with prosodic condition, syllable position, and trial as the within-subject factors. Age group (adult, 8-year-old, 4-year-old) was the between-subject factor. It is to be noted that in order to make this analysis possible, a balanced model, in which each speaker contributes an equal number of repetitions, is necessary. Thus, five repetitions were randomly selected among the set of tokens, in each prosodic condition, and for each speaker.

Morevover, because of the important between-speaker variability often reported in previous studies (Erickson, 1998; Tabain, 2003)**,** individual subject's behaviour are described. Finally, linear regression analyses were carried out on articulatory (as the independent variable) and acoustic values (as the dependent variable).

RESULTS

Articulatory data

The following section groups the values of the maximal lip area, peak velocity of the opening gesture and duration of the opening gesture, in both syllable position (1^{st} or 2^{nd} syllable in words) and prosodic condition (neutral or focus).

Maximal lip area

Mean values and standard deviations of the maximal lip area reached during the production of [a] are plotted in Figure XXX.3, for each speaker. Detailed statistical results are provided in Table XXX.2 (repeated-measures ANOVAs). It is striking to observe that the global values of maximal lip area do not show a tendency to decrease with speaker age. If one considers the size of the articulators for children relative to adults, one could expect that the maximal lip area should be reduced for children. However, according to the results of repeated-measures ANOVAs, the differences related to age as a main effect do not reach significance ($F(2,7)=0.15$; $p>0.05$). This result might be considered as evidence that children hyperarticulated the target syllables more than adults. Concerning the main effect of position, the lip area is significantly larger in the second position than in the first position ($F(1,7)=21.19$; $p<0.05$). This is consistent with the final lengthening effect observed in French. If the second syllable is longer, more time is available for the lips to reach their open configuration. The interaction of position and condition is significant ($F(1,7)=18.15$; $p<0.05$), and planned comparisons reveal that in the first position, lip area is larger for [a] when articulated under contrastive focus, compared to the neutral condition, which is consistent with the perceptual analysis reported previously. No significant difference is observed for the second position. However, as suggested by the significant effect of the interaction of position, condition, and age ($F(2,7)=4.68$; $p<0.05$), a developmental trend is observed. According to the results of planned comparisons, the increased value of lip area in contrastive focus, compared to the neutral condition, in the first syllable position, is significant only for the adult group. Thus, the pattern of variation in lip area shows that all speakers produce larger lip areas in the second position, compared to the first position, and that in the first position, only adult speakers increase the lip area values under contrastive focus, compared to the neutral condition.

Figure XXX.3 also reveals important between-speaker variations. Three adult speakers (CB, PH, and LA) produce [a] with a smaller lip area under focus, compared to the neutral condition, in the second position. Even though this did not reach significance for the adult group, it can be seen as a strategy exploited to enhance the prominence of the first syllable, similarly to the post-focal deaccenting observed in the acoustic domain (as explained previously). Recall that the chosen focus sequences are all perceived as bearing acoustic prominence on the initial syllable. No such pattern is observed for children. One 4-year-old speaker (MA) produces very broad lip areas in the second position, although the standard deviation values are rather large. Concerning the effects of condition, one child out of six (CA, 4) produces much larger values of lip area for both syllables, under contrastive focus, compared to the neutral condition. In other words, this speaker produces decreased values of lip area in the neutral condition compared to the focused condition. The maximum magnitude of lip area for this child is comparable to the mean maximum achieved for the other children, but the minimum (reached in the neutral condition) is lower than the mean for the other children (especially a1 in neutral condition). Therefore, this speaker seems to better exploit the lower range of possible mouth opening for the /a/ vowel.

Peak velocity of the opening gesture

Larger lip area may not always characterize focused syllables. When speaking rate increases, independent of the degree of focus, the amount of time devoted to opening the lips decreases, and the movement may be reduced. Peak velocity can be a better correlate of focus (De Jong, 1995; Kelso *et al.*, 1985; Lœvenbruck, 1996). Our prediction was that higher peak velocities should mark focused syllables. The mean values of the peak velocity of the opening gesture from [b] to [a] are plotted in Figure XXX.4. According to this figure, the peak velocities show a tendency to decrease with decreasing age, except for the 4-year-old speaker CA. However, this global tendency is not significant, based on the results of the repeated-measures ANOVA, reported in Table XXX.2. By contrast, the condition factor has a significant effect on peak velocity ($F(1,7)=9.36$; $p<0.05$), opening movements being faster under contrastive focus than in neutral condition. The interaction between position and condition is also significant ($F(1,7)=25.96$; $p<0.05$), and planned comparisons reveal that peak velocity is increased in focus compared to neutral condition for the first position only. No effect of interaction with age is found, suggesting that all speakers produce the same pattern.

However, within-speaker variability is observed in Figure XXX.4. Two adult speakers, (CB and PH) decrease the values of peak velocity under focus, for the second syllable only. These adults produce a velocity profile in line with the strategy observed previously for lip area: under contrastive focus, the velocity of the opening gesture is increased for the first syllable, whereas the values of the peak velocity for the second position decrease or do not differ according to the prosodic condition. As regards the young speakers, the 4-year-old CA produces much faster opening gestures under contrastive focus than in neutral condition. The very large difference between focus and neutral for this child may be responsible for the lack of significant effect of the factor Age, in the repeated-measures ANOVAs reported in Table XXX.2. This difference is indeed comparable to the difference in velocity produced by the adult speaker AR. Furthermore, unlike what was observed on lip area values, the peak velocities for CA are not just lower for unfocused syllables. For focused syllables they are much higher than they are in the other children. As mentioned above, peak velocity is considered as a good correlate of focus and has been related to articulatory effort or precision. The fact that this speaker is able to increase the velocity of the opening gesture to a degree comparable to the adults' is another argument for concluding that this child is articulatorily more mature

Duration of the opening gesture

Mean duration values (in msec) of the opening gesture are plotted in Figure XXX.5. The global tendency of increasing mean values of duration with decreasing age is significant, based on the results of repeated measures ANOVA (Table XXX.2) ($F(2,7)=25.60$; $p<0.05$). The significant effect of the position factor, without interaction with age, ($F(1,7)=154.40$; $p<0.05$) reveals that all speakers significantly increase the duration of this gesture for the second [a], compared to the first [a]. This pattern is presumably due to accentual phrase final lengthening in both conditions, as explained above. No significant effect of condition is observed. Thus, unlike the pattern revealed by the maximal lip area, duration is not always increased for [a] in the first syllable, under contrastive focus, compared to the neutral condition. Interestingly, this pattern contrasts with English, for which is has been shown that under contrastive focus, the opening gesture is longer, faster, and larger (Cho, in press). This result may be an artefact of the relatively low sampling rate (one sample every 20 ms). Another possible interpretation is that increasing the duration of the opening gesture is not necessary to signal prosodic prominence coming from contrastive focus in French. Increasing articulatory magnitude and movement velocity could be the strategy used to reach the more articulated lip configuration, rather than increasing movement duration.

Figure XXX.5 shows that different durational patterns are produced by speakers. Indeed, some speakers reduce the duration of the opening gesture under focus, in the second position (the adult LA and the children JU, CA, and NE). The difference in duration between the focus and neutral conditions for CA and NE, in the second position, is even larger than the difference found for adults. Note however the large token-to-token variability associated with these values for CA and NE, as revealed by the standard deviations. By contrast, the 8-year-old child LR increases the duration of this movement under focus, in the second position.

Acoustic results

These reinforced articulatory gestures, observed for adults, and sometimes for children, have acoustic consequences in the speech signal. It was predicted that a hyperarticulated [a] (increased lip area, increased tongue lowering and backing) should result in a more peripheral (lower and more back) position of this vowel in the acoustic F1/F2 space, i.e. in a higher F1 and a lower F2. This section investigates the effects of position and prosodic condition on the formant patterns and RMS amplitude values of the vowels in [baba].

Formant values

Dispersion ellipses of the formant values (in Hertz), in the F1 vs. F2 space, are plotted in Figure XXX.6, for both conditions and positions. As reported in other studies (Smith & Gartenberg, 1984; Sharkey & Folkins, 1985), children show a larger token-to-token variability than adults, as depicted by the size of the dispersion ellipses. As can be seen from Table XXX.2, age has a significant effect on F1 values, adult speakers producing lower values than 4-year-old and 8-year-old speakers ($F(2,7)=10.91$; $p<0.05$). This pattern is a consequence of the shorter cavities in the child's vocal tract, compared to the adult's vocal tract. Prosodic condition has a significant effect on F1, the low vowel having higher F1 values under contrastive focus than in neutral condition ($F(1,7)=6.19$; $p<0.05$). This contrast is however not significant in the second position, as reveal planned comparisons of the interaction of position and condition ($F(1,7)=6.79$; $p<0.05$). This is in line with the perceived initial prominence in the focus condition. The effect of age, in interaction with position or condition, is not significant, suggesting that all speakers increase the value of F1 for [a], in the first position. This result is particularly interesting since it does not exactly reflect the articulatory effects of prosodic condition. Indeed, it has been shown, in a previous section, that lip area was not increased in focus, compared to the neutral condition, for the 8-year-old and 4-year-old groups. An inspection of Figure XXX.6 however shows between-speaker variability in the acoustic space, suggesting that some children do not contrast prosodic conditions along the F1 dimension as much as adults do. Indeed, the dispersion ellipses for the first position are much more peripheral under focus (labelled “a1_f”) than in neutral condition (labelled “a1_n”) for three adults out of four (CB, PH, and LA) and one child out of six (the 4-year-old CA). In line with what has been observed in the articulatory domain, speaker CA seems to better exploit the F1 dimension compared to the other children. Child CA uses a wider range of F1 values than any other child and has an adult-like use of F1 differences to convey differences in focus. Among the five remaining children, speaker NE (4-year-old) produces a pattern similar to CA (and to the adult AR), but to a lesser extent. This child was not characterized by increased lip area on the first [a] under focus. Thus, other strategies than increasing lip area are likely exploited to increase F1. The four remaining children produce ellipses associated with “a1_n” and “a1_f” that greatly overlap. The relationships between articulatory and acoustic parameters will be discussed in a further section.

As concerns the variation of F2, Table XXX.2 shows that this formant is significantly higher with increasing age ($F(2,7)=6.75$; $p<0.05$). This effect can be ascribed to the shorter cavities of the children, compared to the adults. No significant effect of position and condition are found. Figure XXX.6 shows that, among the adult speakers, CB and PH produce lower values of F2 for [a] in the second position (ellipses labelled “a2_f” and “a2_n”) under focus than in neutral condition, resulting in a further back position in the vowel spaces. This pattern is also found for English by Cho (in press) who reports that under accent, F2 is lower for [a], i.e. further back in the vowel space. Note that F2 values for [a], produced by the 4-year-old MA, are higher in the second position than in the first position.

RMS amplitude

It has been noted that intensity can sometimes be used to signal contrastive focus. In order to compare this acoustic parameter across syllable positions and prosodic conditions, we calculated the RMS amplitude produced for [a] by each speaker. Data are presented in Figure XXX.7, and p values associated with ANOVAs are grouped in Table XXX.2. Based on the results of Table XXX.2, position has a significant effect on RMS values ($F(1,7)=15.87$; $p<0.05$). The interaction of position and age is also significant ($F(2,7)=6.45$; $p<0.05$). Planned comparisons aimed at further exploring this interaction effect revealed that smaller RMS values for the second position, compared to the first position, are found for the adult group only. As concerns prosodic condition, for both syllable positions, [a] is produced with larger RMS values under focus than in neutral condition ($F(1,7)=6.16$; $p<0.05$).

Individual speakers' data, depicted in Figure XXX.6, suggest that increasing RMS values is an optional strategy to signal contrastive focus, in all age groups. As noted for articulatory measurements, child CA produces maximal RMS values which are higher than any other child, and minimal values (in the neutral condition) comparable to those of adults.

Relationships between articulatory and acoustic parameters

According to the predictions stated above, increased articulatory movements should be related to increased F1 values and decreased F2 values for /a/, resulting in a more peripheral position in the acoustic space. A significant increase in lip area for the first syllable was found for the adult group, only under focus compared to the neutral condition. By contrast, in the acoustic domain, increased F1 values are found for all age groups, without any significant interaction of age. It is likely that other articulatory strategies, such as tongue movements, allowed some speakers to produce a more peripheral [a], in the acoustic space. Furthermore, under the effect of position, speakers produce broader lip areas without any effect in the F1 or F2 dimensions (see the

F-values grouped in Table XXX.2). In order to explore the relationship between lip area and acoustic parameters, results of linear regression analyses for each speaker are grouped in Table XXX.3. The correlation coefficients r and the slope are presented only for significant correlations. Lip area values have been regressed against F1, F2, and RMS. It can be seen that significant correlations between maximal lip area and F1 are obtained for all speakers but 2 (the 8-year-old LR and the 4-year-old MA). Even though no developmental pattern can be found in the values of the slope, it is noteworthy that the three adult speakers CB, PH, and LA are associated with the shallower slopes, suggesting that an increase in lip area is related to a smaller increase in F1 for these speakers, compared to the other speakers. Lip area and F2 are significantly correlated for the 4-year-old speaker MA and the adult CB, in line with Cho (in press). Note that RMS is correlated with lip area for two adults (AR and PH) and one 4-year-old child (CA). The slope values for these correlations are however very small. No developmental pattern emerges. Overall, F1 seems to be correlated with lip area in adults and children, even though the strength of the correlation varies across speakers.

DISCUSSION

The results presented so far have shown that the articulatory effects of contrastive focus for /a/ are smaller for 4-year-old and 8-year-old children compared to adults. Indeed, repeated-measures ANOVAs reveal that in the first syllable position, maximal lip area values are larger under contrastive focus that in neutral condition for the adult group only. This result is in line with previous studies showing that children by 4 years of age have not yet mastered the full scale of articulatory contrast related to prosodic prominence. A possible explanation could be that children did not understand the task or signalled contrastive focus using only syntactic extraction.

However, as described in the Method section, prosodic focus was very well signalled by a F0 rise on the first syllable, followed by a typical deaccented post-focus sequence. Thus, children as well as adults used prosody to convey focus and the observed differences can not be ascribed to attention or linguistic abilities. As concerns peak velocity, opening movements from /b/ to /a/ are faster in focus than in neutral condition, for the first syllable. Opening movements are globally longer for children than for adults, and their duration are not significantly related to condition. For all age groups, F1 is higher and RMS values are greater in the first position under focus than in neutral condition. In order to better describe each speaker's behaviour, individuals' data were presented and discussed. First, it was found that for three adult speakers, lip area is wider on the first syllable, as predicted, but lip area decreases in the second syllable, under contrastive focus. This pattern has been interpreted as a strategy used by the adult speakers to enhance the prominence of the first syllable, by hypoarticulating the second syllable, similarly to the post-focal F0 reduction observed in the acoustic domain. This lip area pattern is accompanied by an increased velocity of the opening gesture, and by a reduced duration of the opening gesture (for two adults). The latter result contrasts with English for which it was observed (Cho, in press) that contrastive focus is related to longer opening gestures. Nonetheless, considering lip area and velocity values, these results suggest that adult speakers of French contrast focused and unfocused syllables along the articulatory dimension of lip area (see Lœvenbruck, 1999 for similar findings with tongue gestures and Dohen *et al.* 2004a and b for further results on lip area and jaw opening). Such a pattern was not observed in children, suggesting that they do not exploit the lower range of articulatory contrasts.

The case of the 4-year-old speaker CA is particularly interesting. Unlike other children, this speaker shows focus-induced variations in lip area and opening peak velocity (cf Figures

XXX.3-XXX.4). As mentioned earlier, CA even seems to better exploit the full scale of articulatory contrasts along the lip area dimension. Indeed, the maximum magnitude of lip area for this child is comparable to the mean maximum achieved for the other children, but the minimum (reached in the neutral condition) is much lower than the mean for the other children (especially a1 in neutral condition). Note that the mean values of opening peak velocity are even higher for this child compared to adults, and could be responsible for the lack of significant results as concerns the effect of age in interaction with condition. The fact that this speaker is able to increase the velocity of the opening gesture to a degree comparable to the adults' is another argument for concluding that this child is articulatorily more mature.

In the formant domain, besides a significant effect of age on F1 and F2, no significant developmental trends were found. For all speakers, F1 is significantly higher under focus than in neutral condition, in the first position. These effects are in line with previous works carried out on articulatory and acoustic effects of focus in English (Erickson, 1998; Cho, in press). An inspection of individual strategies however suggested that the acoustic effects of contrastive focus is less consistent in children compared to adults. Indeed, four children produced overlapping dispersion ellipses, in the F1 vs. F2 space, while two 4-year-olds produced a pattern in line with the one produced by the adult speakers. Further experiments are currently in progress to support this finding, since the small number of repetitions (five) and the large variability could have affected the results.

It was also found that for some children, maximal lip area could be increased while F1 was not significantly affected. This dissociation between articulatory and acoustic correlates is not reflected by smaller correlation coefficients between lip area and F1. The correlation coefficients in fact suggest that F1 is as much related to lip area for children as it is for adult speakers. Some children do not show any significant correlation between lip area and F1, whereas some others are characterized by high correlation coefficients and slopes. It is likely that other articulatory strategies than lip area, such as tongue movements, allow the production of a more peripheral /a/, in the acoustic space. Previous work has shown that adult-like control is acquired earlier for the jaw than for the lips (Green *et al.*, 2002), mature control of the tongue being reached even later (Nittrouer, 1993). Thus, whereas adults can generate differentiated jaw, lip, and tongue movements to produce a more peripheral [a], children would not differentiate jaw, lip, and tongue movements to the same extent. According to Erickson, Iskarous, and Whalen (2005), contrastive focus is characterized by independent movements of the tongue and the jaw. Some of the young speakers could not have differentiated jaw, lip, and tongue movements, resulting in an overall reduced effect of prosodic prominence on articulatory and acoustic parameters. This highly complex differentiation phase would be an important step in motor control development (Green *et al.*, 2002). By contrast, some children could be in the process of differentiating jaw, lip, and tongue movements, resulting in a more important effect of prosody on one articulator (the tongue, for instance), compared to the others (the jaw and the lips). The necessity of increasing contrasts between prosodic conditions could thus be seen as a factor leading to fine-tuning of different articulatory movements.

On the basis of the global patterns reported here, a hasty conclusion could be that the smaller effects of focus for children are related to the inability of the children to hyperarticulate in this condition. However, the greater values of lip area depicted in Figure XXX.3 suggest that, in line with Smith and Goffman’s observations (1998), children produce relatively large articulatory movements considering the size of their oral structures, compared to adults. Thus, the observed undifferentiated kinematic patterns according to condition could rather be due to the inability to reduce unfocused syllables (Allen and Hawkins, 1980). Children have generally learned to exploit the higher range of the hyper-hypoarticulation scale, but the ability to reduce the magnitude of articulatory movements (i.e. exploiting the lower range of the scale) comes later in the acquisition of motor control, presumably because it may require fine-tuning of articulation. The differences between children and adults may be even more obvious in more complex sequences. Indeed, according to Smith and Goffman (1998), it might be hypothesized that certain phonetic goals (rather than specific articulator controls) may mature earlier than others (/b/ mastered earlier than /v/). The /baba/ sequence involves the combined action of the lips and jaw and it may be that this sequence is acquired earlier, relative to sequences involving tongue movements (such as /tata/) or tightly-controlled lip movements (such as /vava/). Further articulatory and acoustic analyses are currently in progress in order to further investigate this hypothesis.

CONCLUSION

This paper dealt with the articulatory and acoustic correlates of contrastive focus in three 4-year-old and three 8-year-old children, and four adult French speakers. Using video recordings, lip area and its first derivative were studied during the production of the [baba] sequence in neutral condition and under contrastive focus. Results show that the effects of contrastive focus on lip area are less important for children compared to adults. These results suggest that children at age 8 do not contrast the vowels along the hypo-hyperarticulation dimension according to the prosodic condition, as adults do. Learning the articulatory correlates of prosodic contrasts could involve the integration of multiple articulatory movements in order to better exploit the acoustic space.

ACKNOWLEDGEMENTS

This work was supported by a grant from the Social Sciences and Humanities Research Council of Canada and a grant from the Fonds Québécois de Recherche sur la Société et la Culture. We thank the reviewers for constructive comments on an earlier version of this paper. This work benefited from fruitful discussions with Christian Abry and Pascal Perrier.

REFERENCES

Allen, G. D., and Hawkins, S. (1980): “Phonological rhythm: Definition and development”, in Yeni-Komshian, G., Kavanagh, J. F. and Ferguson, C. A. (eds.), *Child Phonology Vol. 1: Production*. New York: Academic Press, 227-256.

Audouy, M. (2000): *Traitement d’images video pour la capture des mouvements labiaux*, Final engineering report, Institut National Polytechnique de Grenoble, 2000.

Bartels, C. and Kingston, J. (1994): “ Salient Pitch Cues in the Perception of Contrastive Focus”, in Boach, P. and Van der Sandt, R. (eds), *Focus & Natural Language Processing*, Proceedings of J. Sem. conference on Focus, IBM Working Papers, TR-80, 94-106.

Beckman, M. E., Edwards, J., and Fletcher, J. (1992): “Prosodic structure and tempo in a sonority model of articulatory dynamics”, In Docherty, G. J. and Ladd, D. R. (eds), *Papers in Laboratory Phonology II: Segment, Gesture, Prosody*, Cambridge University Press, 68-86.

Cho. T. (2002): *The Effects of Prosody on Articulation in English*, New York: Routledge.

Cho, T. (in press):”Prosodic strengthening and featural enhancement: Evidence from acoustic and articulatory realizations of /ɑ, i/ in English, *Journal of the Acoustical Society of America*.

Clech-Darbon, A., Rebuschi, G., and Rialland, A. (1999): “Are there Cleft Sentences in French?” in Tuller, L. and Rebuschi, D. (eds), *The Grammar of Focus*, Amsterdam, Benjamins, 83-118.

Connaghan, K. P., Moore, C. A., Reilly, K. J., Almand, K. B., and Steeve, R. W. (2001): “Acoustic and physiologic correlates of stress production across systems”, *Poster presented at the American Speech-Language-Hearing Association*, New Orleans, L. A.

Dahan, D., & Bernard, J.-M. (1996): “Interspeaker variability in emphatic accent production in French, *Language and Speech*, 39, 341-374.

De Jong, K. (1995): “The supraglottal articulation of prominence in English : linguistic stress as localized hyperarticulation” *Journal of the Acoustical Society of America*, 97, 491-504.

Delais-Roussarie, E., Rialland, A., Doetjes, J., Marandin, J. M. (2002): “The Prosody of Post Focus Sequences in French”, in *Proceedings of Speech Prosody 2002*, Aix-en-Provence.

Di Cristo, A. (2000): “Vers une modélisation de l’accentuation du français (deuxième partie)”, *Journal of French Language Studies*, 10, 27-44.

Di Cristo, A. (1998): “Intonation in French”, in Hirst, D. & Di Cristo, A. (eds), *Intonation Systems: A Survey of Twenty Languages*, Cambridge, Cambridge University Press, 195-218.

Dohen M. & Loevenbruck H. (2004): Pre-focal Rephrasing, Focal Enhancement and Post-focal Deaccentuation in French. *Proceedings of International Conference on Spoken Language Processing* (ICSLP), Jeju island (Korea), 4-8 October 2004, Vol. 1, 785-788.

Dohen M., Lœvenbruck H., Cathiard M.-A. & Schwartz J.-L. (2004a).): Visual perception of contrastive focus in reiterant French speech. *Speech Communication,* 44, 155-172.

Dohen M., Lœvenbruck H., Cathiard M.-A. & Schwartz J.-L. (2004b): Identification of the possible visible correlates of contrastive focus in French. *Proceedings of Speech Prosody* 2004, Nara (Japon), 23-26 March 2004, 73-76.

Erickson, D. (1998): "Effects of Contrastive Emphasis on Jaw Opening", *Phonetica,* 55, 147-169.

Erickson, D. (2002): "Articulation of Extreme Formant Patterns for Emphasized Vowels", *Phonetica,* 59, 134-149.

Erickson, D. , Iskarous, K., and Whalen, D. (2005): "Production of hyperarticulated vs. contrastively accented vowels", *A Century of Experimental Phonetics: Its History and Development*, Grenoble, Institut de la Communication Parlée.

Goffman, L. and Malin, C. (1999): "Metrical Effects on Speech Movements in Children and Adults", *Journal of Speech, Language, and Hearing Research,* 42 (4), 1003-1015.

Green, J. R., Moore, C. A., Higashikawa, M., and Steeve, R. W. (2000): "The Physiologic Development of Speech Motor Control : Lip and Jaw Coordination", *Journal of Speech, Language and Hearing Research,* 43, 239-255.

Green, J.R., Moore, C.A., and Reilly, K.J. (2002): "The sequential development of lip and jaw control for speech", *Journal of Speech, Language, and Hearing Research,* 45, 66-79.

Hornby, P. A. and Hass, W. A. (1970): "Use of contrastive stress by preschool children", *Journal of Speech and Hearing Research*, 13, 395-399.

Jun, S.-A. and Fougeron, C. (2000): "A phonological model of French intonation", in Botinis, A. (ed), *Intonation: Analysis, Modelling and Technology*, KAP, Dordrecht, 209-242.

Jun, S.-A. and Fougeron, C. (2002): "Realizations of accentual phrase in French intonation", Probus, 14, 147-172.

Kehoe, M., Stoel-Gammon, C., and Buder, E. H. (1995): “Acoustic Correlates of Stress in Young Children’s Speech”, *Journal of Speech, Language, and Hearing Research*, 38, 338-350.

Kelso J.A.S, Vatikiotis-Bateson E, Saltzman E.L. & Kay B. (1985): A qualitative dynamic analysis of reiterant speech production: phase portraits, kinematics, and dynamic modeling’. *J. Acous. Soc. Am.*, 77 (1), 266-280.

Kent, R. D. & Forner, L. L. (1980): “Speech segment durations in sentence recitations by children and adults”, *Journal of Phonetics*, 8, 157-168.

Konopscynsky, G. (1986): *Du prélangage au langage : acquisition de la structuration prosodique,* Thesis, University of Strasbourg.

Lacheret-Dufour, A. and Beaugendre, F. (1999): *La prosodie du français*, Paris, CNRS Langage.

Ladd, R. D. (1996): *Intonational Phonology*, Cambridge, Cambridge University Press.

Lœvenbruck H. (1996): Pistes pour le contrôle d’un robot parlant capable de réduction vocalique. *PhD dissertation, INPG, Grenoble, France.*

Lœvenbruck, H. (1999): “An investigation of articulatory correlates of the accentual phrase in French”, *Int. Congr. Phon. Sc.*, San Francisco, Vol. 1, 667-670.

Max, L. and Onghena, P. (1999): “Some Issues in the Statistical Analysis of Completely Randomized and Repeated Measures Designs for Speech, Language, and Hearing Research”, *Journal of Speech, Language, and Hearing Research*, 42, 261-270.

Nittrouer, S. (1993): “The Emergence of Mature Gestural Patterns is Not Uniform : Evidence from an Acoustic Study”, *Journal of Speech and Hearing Research*, 36, 959-972.

Pasdeloup, V. (1990): *Un modèle de règles rythmiques appliqué à la synthèse de la parole*, Ph. D. thesis, Aix-en-Provence, Laboratoire Parole et Langage.

Pierrehumbert, J., and Hirshberg, J. (1990): “The meaning of intonational contours in discourse”, in Cohen, P., Morgan, J., and Pollack, M. (eds), *Intentions in Communication*, Cambridge, MIT Press.

Rossi, M. (1999): “La focalisation”, in *L’intonation, le système du français: description et modélisation*, Ophrys, 116-128.

Selkirk, E. O. (1984): “The grammar of intonation”, in Selkirk, E. O. (ed), *Phonology and Syntax: The Relation between Sound and Structure*, Cambridge, MIT Press, 197-296.

Sharkey, J. & Folkins, J. W. (1985): “Variability of lip and jaw movements in children and adults: implications for the development of speech motor control”, *Journal of Speech and Hearing Research*, 28, 8-15.

Smith, B. L. & Gartenberg, T. E. (1984): “Initial observations concerning developmental characteristics of labio-mandibular kinematics”, *Journal of the Acoustical Society of America*, 75, 1599-1605.

Smith, A. & Goffman, L. (1998) “Stability and patterning of speech movement sequences in children and adults” *Journal of Speech, Language, and Hearing Research,* 41, 18-30.

Tabain, M. (2003): “Effects of prosodic boundary on /aC/ sequences: articulatory results”, *Journal of the Acoustical Society of America* ,113, 2834-2849.

Touati, P. (1987): *Structures prosodiques du suédois et du français,* Working Paper 21, Lund University Press.

Vihman, M. M. (1996): *Phonological Development,* Cambridge/Oxford, Blackwell Publishers.

Welby, P. (2003): *The slaying of Lady Mondegreen, being a study of the association and alignment of French intonational rises and their role in speech segmentation*, Ph. D. thesis, Ohio, The Ohio State University.

Figure captions

Figure XXX.1: Example of a video image with the lip area represented by the zone shaded by diagonal lines.

Figure XXX.2: Labeling points on the [baba] sequence. Upper panel: Acoustic signal; Middle panel: Lip area (S) (in cm^2); lower panel: velocity (first derivative of the S values). These events correspond to the zero-crossing, minimal and maximal values of the two signals. Note that due to the sampling rate (50 Hz) of the video signal, one sample occurs every 20 ms.

Figure XXX.3: Mean values and standard deviations of the maximal lip area for the 12 speakers, in neutral condition (light bar) and under contrastive focus (shadow bar), for the first vowel (labelled “a1”) and the second vowel (labelled “a2”) in [baba].

Figure XXX.4: Mean values and standard deviations of the peak velocity of the opening gesture (first derivative of the lip area) for the 12 speakers, in neutral condition (light bar) and under contrastive focus (shadow bar), for the first vowel (labelled “a1”) and the second vowel (labelled “a2”) in [baba].

Figure XXX.5: Mean duration values and standard deviation for the opening gesture for the 12 speakers, in neutral condition (light bar) and under contrastive focus (shadow bar), for the first vowel (labelled “a1”) and the second vowel (labelled “a2”) in [baba].

Figure XXX.6: Dispersion ellipses, in the F1 vs. F2 space, enclosing ±1.5 standard deviations around the mean, for the 12 speakers (thin dotted line: first vowel in neutral condition (a1_n); thin solid line: second vowel in neutral condition (a2_n); thick dotted line: first vowel in contrastive focus (a1_f); thick solid line: second vowel in contrastive focus (a2_f)).

Figure XXX.7: Mean values and standard deviations of the RMS amplitude, for the 12 speakers, in neutral condition (light bar) and under contrastive focus (shadow bar), for the first vowel (labelled “a1”) and the second vowel (labelled “a2”) in [baba].

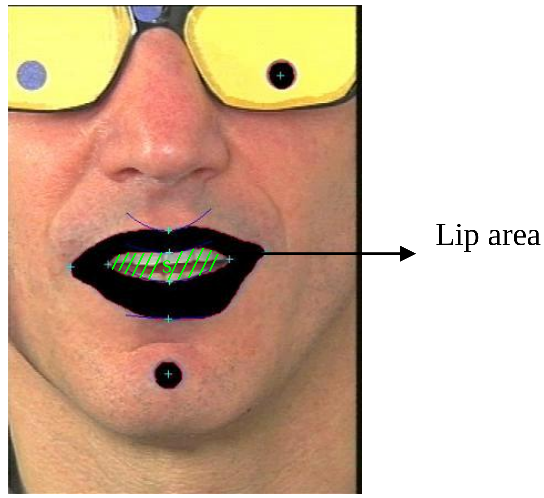


Figure XXX.1

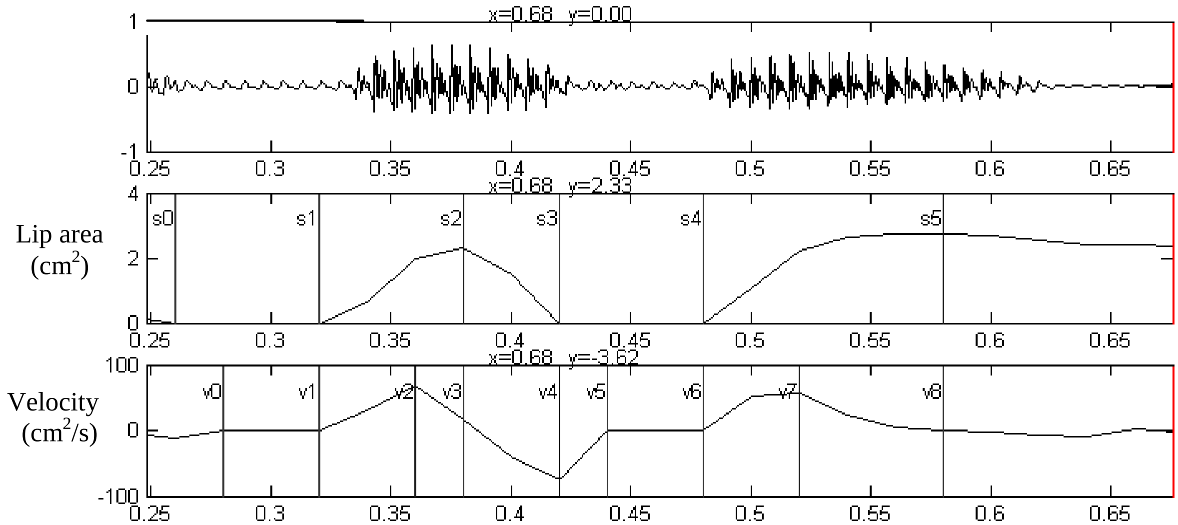

x=0.68 y=0.00
x=0.68 y=2.33
x=0.68 y=-3.62
Lip area
(cm²)
Velocity
(cm²/s)
s0
s1
s2
s3
s4
s5
v0
v1
v2
v3
v4
v5
v6
v7
v8

Figure XXX.2

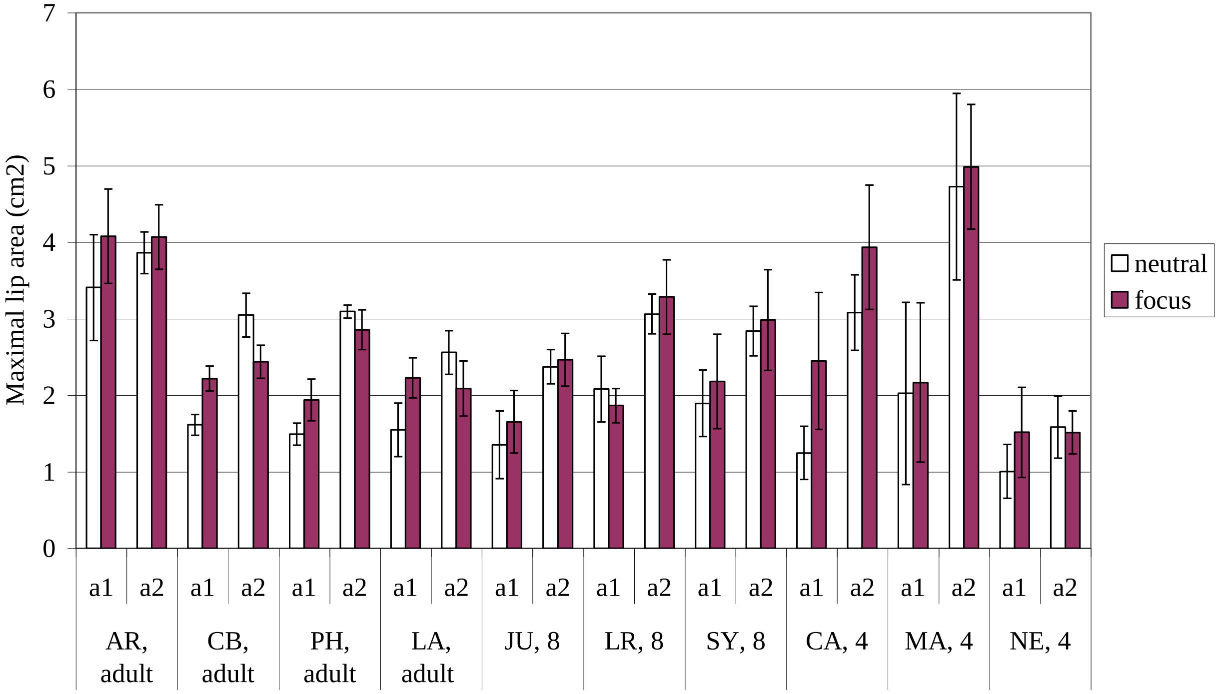
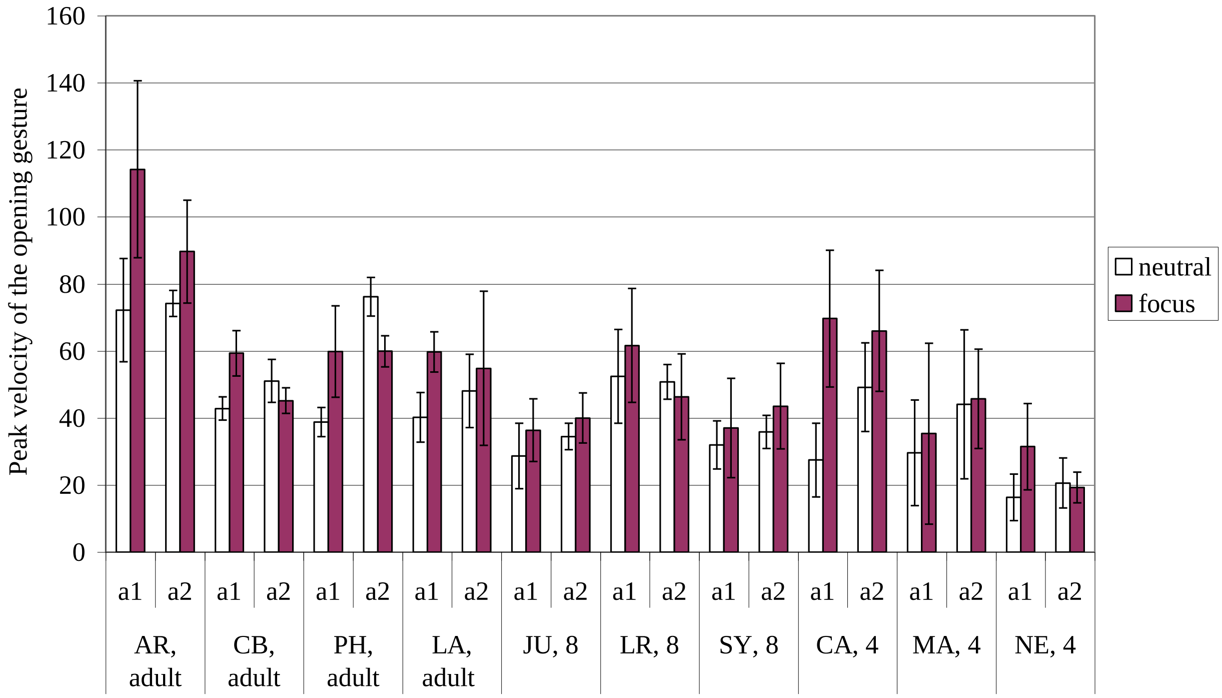
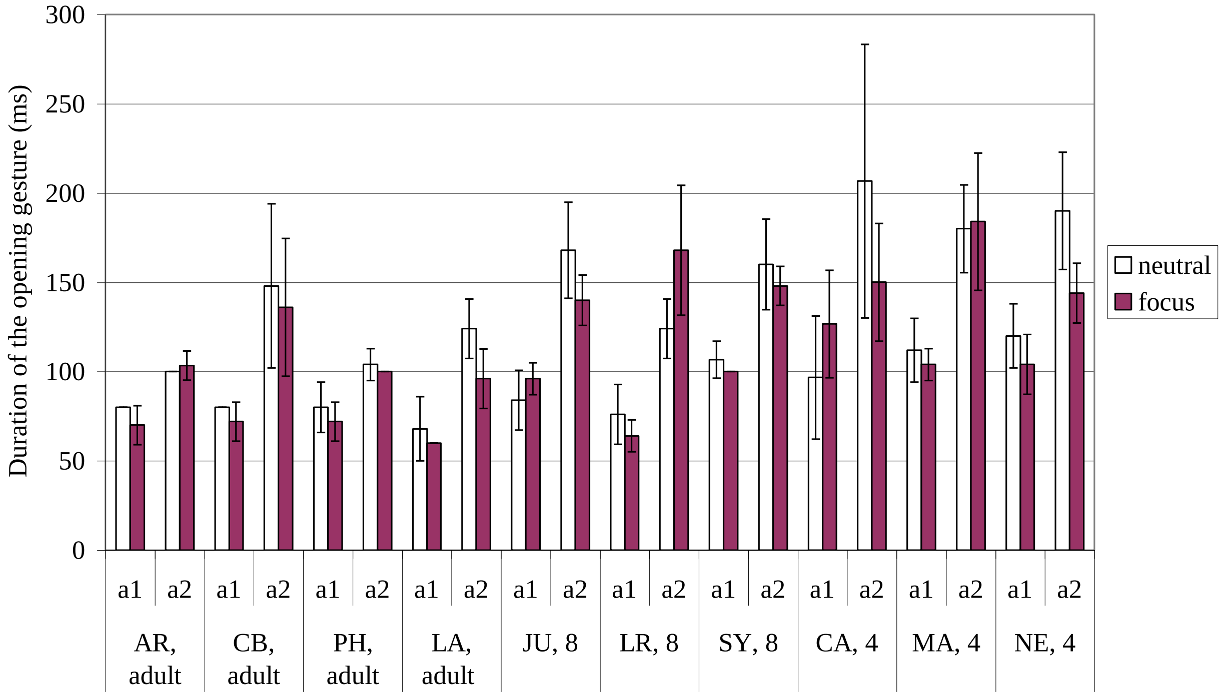

Maximal lip area (cm2)
7
6
5
4
3
2
1
0
neutral
focus
a1
a2
AR, adult
CB, adult
PH, adult
LA, adult
JU, 8
LR, 8
SY, 8
CA, 4
MA, 4
NE, 4

Figure XXX.3



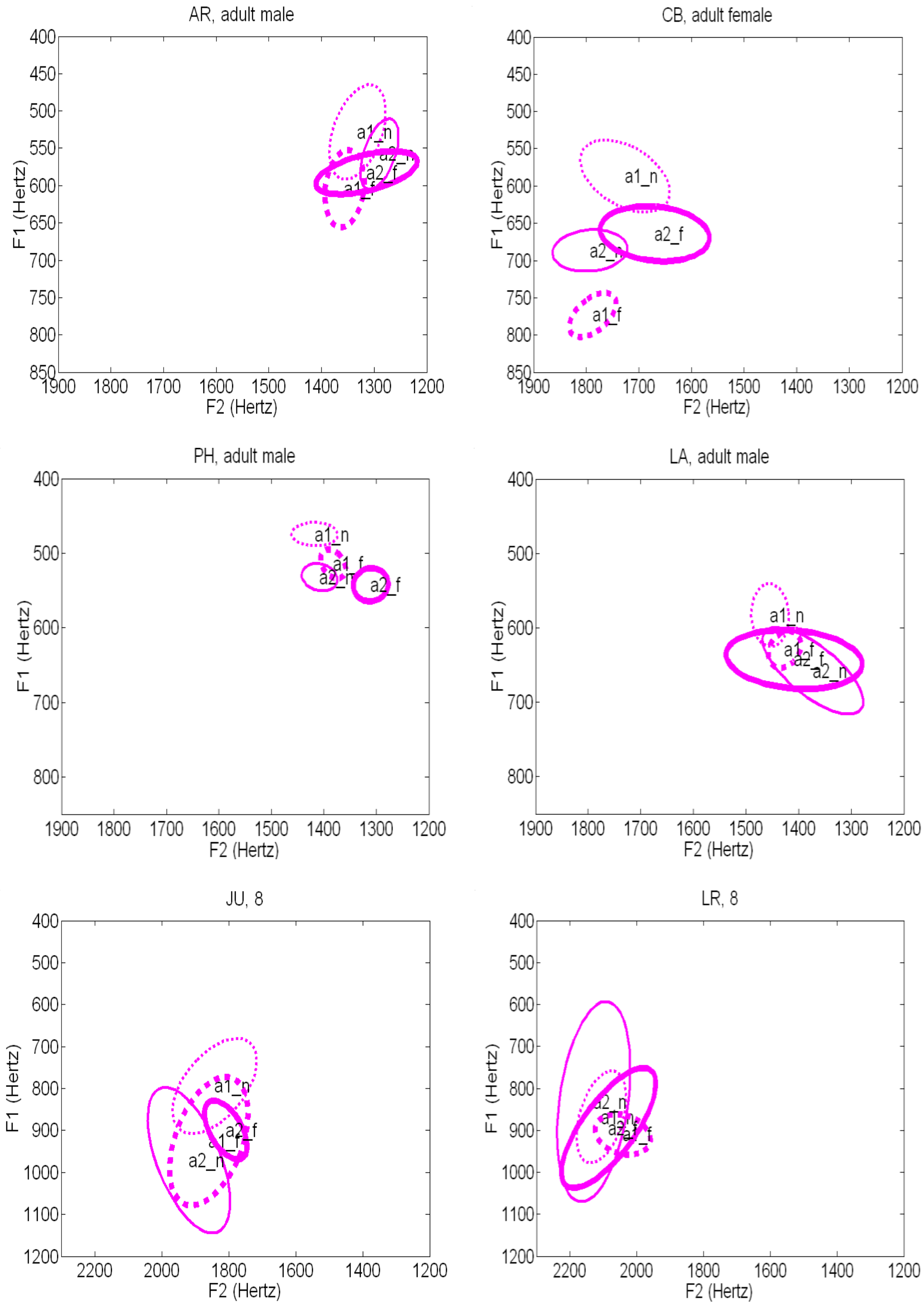
Peak velocity of the opening gesture
160
140
120
100
80
60
40
20
0
neutral
focus
a1
a2
AR, adult
CB, adult
PH, adult
LA, adult
JU, 8
LR, 8
SY, 8
CA, 4
MA, 4
NE, 4

Figure XXX.4



Duration of the opening gesture (ms)
300
250
200
150
100
50
0
neutral
focus
a1
a2
AR, adult
CB, adult
PH, adult
LA, adult
JU, 8
LR, 8
SY, 8
CA, 4
MA, 4
NE, 4

Figure XXX.5


AR, adult male
CB, adult female
PH, adult male
LA, adult male
JU, 8
LR, 8
F1 (Hertz)
F2 (Hertz)
a1_n
a1_f
a2_n
a2_f

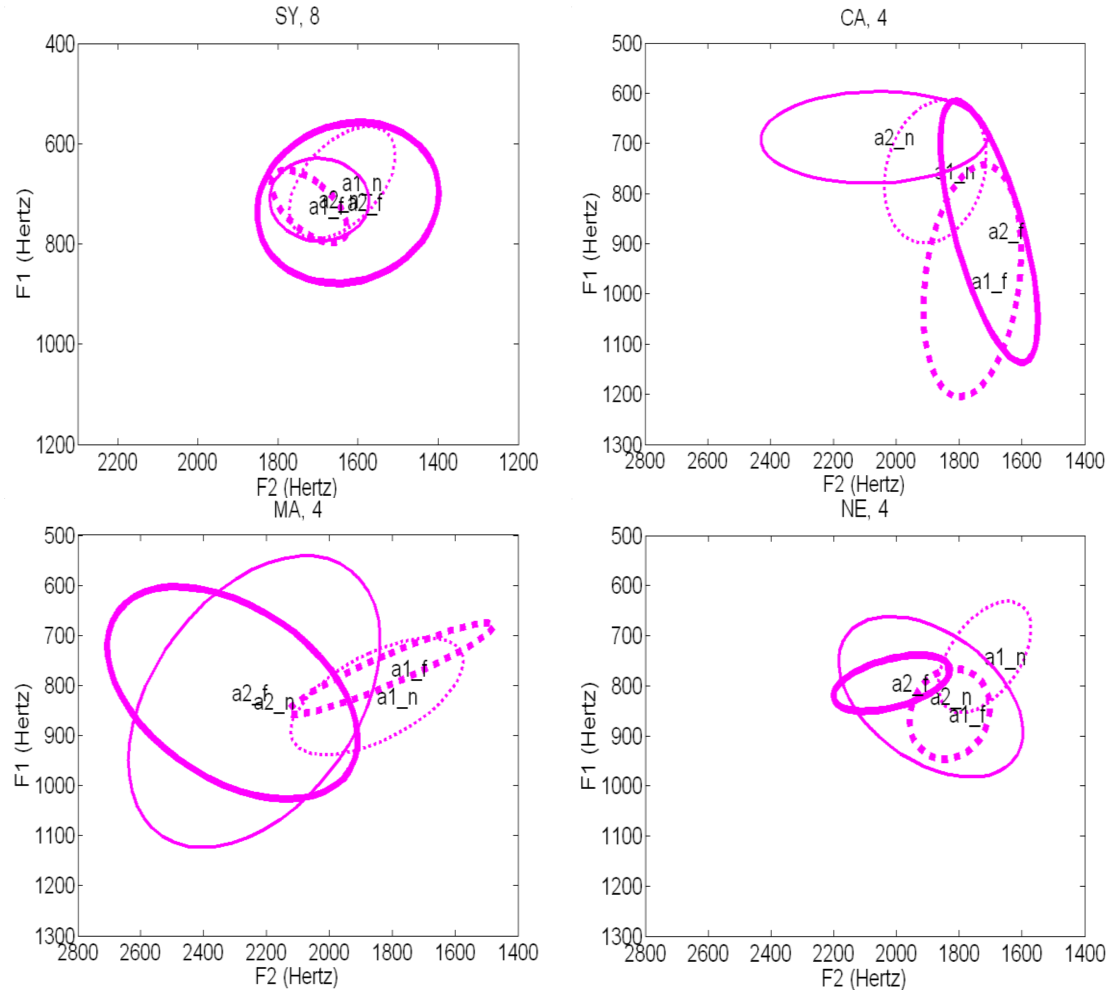
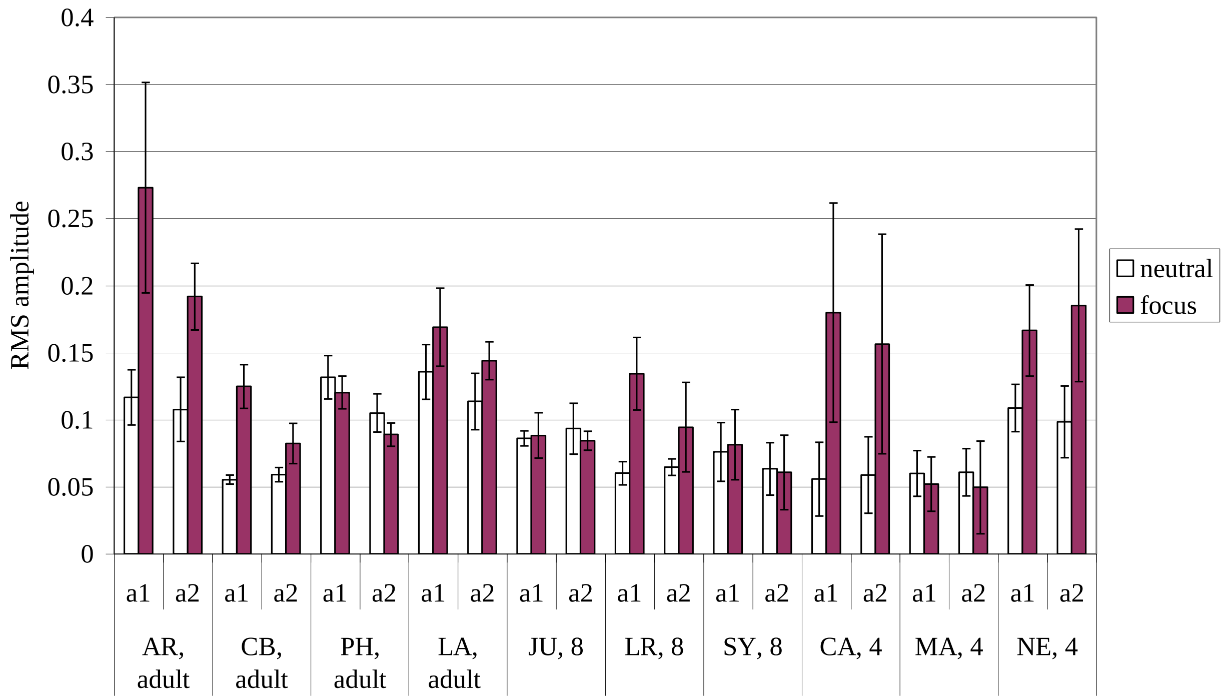

SY, 8
F1 (Hertz)
F2 (Hertz)
a1_n
a2_f
a1_f
a2_n
CA, 4
a2_n
a1_n
a2_f
a1_f
MA, 4
a2_f
a2_n
a1_f
a1_n
NE, 4
a1_n
a2_f
a2_n
a1_f

Figure XXX.6



0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
RMS amplitude
neutral
focus
a1
a2
AR, adult
CB, adult
PH, adult
LA, adult
JU, 8
LR, 8
SY, 8
CA, 4
MA, 4
NE, 4

Figure XXX.7

Table XXX.1 : Number of sentences produced by each speaker, in each prosodic condition (neutral and under contrastive focus), and for each perceived accent pattern. Syllables in capital letters are those perceived prominent by at least four judges among five.

	Neutral	**Contrastive focus**		
	baBA	**BAba**	**baBA**	**Total**
AR, adult	11	6	3	**9**
CB, adult	17	11		**11**
PH, adult	12	9		**9**
LA, adult	12	8		**8**
JU, 8	5	5		**5**
LR, 8	9	5		**5**
SY, 8	6	5	1	**6**
CA, 4	6	6		**6**
MA, 4	5	5		**5**
NE, 4	6	5	1	**6**

Table XXX.2: F values of repeated measures ANOVAs, with condition (neutral or focus), position (first or second syllable), and trial (one of 5 trials) as the within-subject factors. The between subject variable was age. For all speakers, five tokens were randomly selected in each prosodic condition, in order to obtain a balanced model and to make the analysis possible. Values in bold characters and followed by * are those for which $p<0.05$.

Factor	Maximal lip area	Peak vel. open. gesture	Duration of the open. gesture	F1 (Hz)	F2 (Hz)	RMS ampl.
Age F(2,7)	0.15	2.93	**25.60***	**10.91***	**6.75***	1.38
Pos F(1,7)	**21.19***	0.58	**154.40***	1.71	2.84	**15.87***
Cond F(1,7)	3.95	**9.36***	5.60	**6.19***	0.67	**6.16***
Pos*Cond F(1,7)	**18.15***	**25.96***	1.00	**6.79***	3.57	2.94
Pos*Age F(2,7)	0.93	0.23	3.10	0.31	5.25	**6.45***
Cond*Age F(2,7)	0.87	0.51	1.30	0.46	0.02	0.62
Pos*Cond *Age F(2,7)	**4.68***	4.05	1.30	0.23	0.11	1.37

Table XXX.3: Correlation coefficient (r) and slopes obtained from linear regression analyses between maximal lip area and F1, maximal lip area and F2, and maximal lip area and RMS values. Analyses were performed for each speaker, across position and condition. Only significant correlations are shown.

Speaker	**Max. lip area vs. F1**		**Max. lip area vs. F2**		**Max. lip area vs. RMS**	
	r	**slope**	**r**	**slope**	**r**	**slope**
AR, ad.	0.84	71.28	-	-	0.60	0.09
CB, ad.	0.48	52.99	0.30	40.53	-	-
PH, ad.	0.83	40.93	-	-	0.67	-0.02
LA, ad.	0.62	55.82	-	-	-	-
JU, 8	0.60	103.91	-	-	-	-
LR, 8	-	-	-	-	-	-
SY, 8	0.62	66.10	-	-	-	-
CA, 4	0.46	61.93	-	-	0.58	0.04
MA, 4	-	-	0.69	75.88	-	-
NE, 4	0.60	106.17	-	-	-	-