Invariance and variability in the production of the height feature in French vowels

Lucie Ménard, Jean-Luc Schwartz, Jérôme Aubin

Invariance and variability in the production of the height feature in French vowels

Lucie Ménard[a], Jean-Luc Schwartz[b], and Jérôme Aubin[a]

[a] Département de linguistique et de didactique des langues

Université du Québec à Montréal, Montreal

[b] Institut de la Communication Parlée

INPG/Université Stendhal, Grenoble

Running head: Height feature in French

Abstract

This paper investigates the organization of the vowel space in French speakers. Speakers from 4 years of age to adulthood were recorded in order to generate significant between-speaker variability. Each speaker produced repetitions of the ten French oral vowels /i y u e ø o ɛ œ ɔ a/. Acoustic analyses show that despite considerable between-speaker variability in the relative positions of the vowels within the vowel space, speakers tend to produce vowels along a given height degree with a stable F1 value, depending on the speaker, but independently of place of articulation and roundedness. Simulations with the Variable Linear Articulatory Model (VLAM) show that a stable F1 value is basically related to stable tongue heights. The results are discussed in the framework of the *Perception-for-Action Control* theory (PACT), in which speech units are considered as gestures shaped by perceptual processes.

1. Introduction

In recent decades, proposals for a substance-based account of the vowel systems of the world's languages have provided insights into the articulatory, acoustic, and perceptual constraints shaping phonological inventories. Lindblom's Dispersion Theory (DT; Lindblom, 1986) states that vowel systems are shaped by a criterion of sufficient perceptual contrasts. In this view, a five-vowel system like /i e a o u/ is favored over a system like /i e ɛ a o/ because of the greater perceptual dispersion and distinctiveness of the former compared to the latter. This relational constraint differs from Stevens' (1989) local constraints, proposed in Quantal Theory (QT). According to QT, preferred phonemes result from the non-linearity of the articulatory and acoustic spaces. Quantal regions of the articulatory-acoustic space are those for which the acoustic pattern is relatively insensitive to variation in articulatory settings.

Inspired by the DT and QT, the Dispersion-Focalization Theory of vowel systems (DFT; Schwartz et al., 1997) assumes that vowel systems are shaped by both global structural dispersion constraints, aimed at maximizing the acoustic distance between vowels (Lindblom, 1996), and local focalization constraints that favor vowels for which two adjacent formants are close together. Focalization has recently been claimed to be important for infant and adult vowel perception, as demonstrated by the asymmetry effect in vowel discrimination (Polka and Bohn, 2003; Schwartz et al., 2005). Furthermore, Ménard et al. (2004, to appear) showed that focalization seems to be part of the speaker's

task in French, sometimes even—in the course of development—at the cost of intelligibility.

Despite the robustness of the dispersion and focalization constraints, the DFT fails to explain why vowel systems with fewer than nine vowels generally do not use secondary features (such as duration and nasality). Rather, the preponderant vowel systems in such languages make use of the three primary features of height, place of articulation, and roundedness, combined according to a principle of sufficient perceptual contrast (for five vowels: /i e a o u/ rather than /ə ɞ ɵ ɜ ɐ/ or /i ẽ a o: u/). Ohala (1979) refers to this pattern as the *Maximum Utilization of the Available Features* (MUAF) principle. According to the MUAF, when a new feature is added to the system, it tends to be systematically combined with available features.

In an integrated theory of speech perception, the *Perception-for-Action Control Theory* (PACT), Schwartz et al. (2002, 2006) propose that speech units are gestures shaped by multisensory perceptual mechanisms. In this view, speech units are neither purely motor in nature (as in Motor Theory; Liberman and Mattingly, 1985) nor purely auditory (cf. Nearey, 1997), but rather emerge from perceptuo-motor processes. Several experiments have provided evidence that acoustic parameters and articulatory knowledge are both part of the speech goal (Perrier, 2005; Schwartz et al., 2006). Within the PACT, speech perception allows a listener to follow and recover the speaker's vocalizations and to control his or her own gestures. Vowel systems are organized according to perceptual distinctiveness and focalization constraints (as in the DFT), combined with a principle of

articulatory regularization (see also the Lexical Recalibration Model; Lindblom, 1998). According to this principle, instead of maximizing the use of features, systems maximize the use of available articulatory controls. Once height and place of articulation are controlled (for instance, in a four-vowel system /i e a u/), when a new unit is added, it will be combined with the available controls (in a five-vowel system /i e a o u/, where both /e/ and /o/ are mid-high vowels) instead of adding a new control (for instance, nasalization, as in /i e a õ u/). Such articulatorily regularized systems are said to be easier to learn.

The PACT therefore replaces the Maximal Use of Available Features principle proposed by Ohala by a Maximal Use of Available Controls (MUAC) principle. Since controls are different from features, the MUAC principle should in some cases make different predictions from the MUAF. This may well be the case, in light of a specific pattern discovered in the French vowel system by Neagu (1997). In a study of 12 French speakers (six male and six female), Neagu (1997) compared the values of F1 for similar height degrees across rounding and place of articulation. Three height degrees were analyzed: high vowels (/i y u/), mid-high vowels (/e ø o/), and mid-low vowels (/ɛ œ ɔ/). For each speaker, the difference in F1 between the low vowel (/a/) and the high vowels (/i y u/) was calculated, to obtain the range of F1 exploited by that speaker. For each mid-high and mid-low vowel, and for each speaker, the value of F1 relative to the range of F1 exploited in the vowel space was determined. The results revealed that, despite a speaker-specific distribution of height degrees along the F1 dimension, speakers showed very little variability in F1 value within a given height degree, across place of articulation and

rounding. This seems to show that a given phonetic feature (here, height) is realized idiosyncratically by implementing a speaker-specific control of the F1 value. Once mastered by the speaker, this control is combined with other controls (in this case, for implementing place and rounding contrasts) to provide a speaker-specific implementation of the whole system, with a speaker-dependent set of F1 values, but a rather stable organization of the vowels in the system around these specific F1 values.

In this paper, we further explore Neagu's finding, namely the tendency to produce vowels of similar height with a stable F1 value, independently of place of articulation or roundedness. To this end, French speakers from different dialect regions (Canadian French and Continental French) spanning different age groups (from 4 years of age to adulthood) are analyzed, to determine to what extent vowel spaces are shaped by perceptual distinctiveness constraints (cf. the DFT) and articulatory regularization principles such as the PACT. The great between-speaker variability of our corpus (Ménard, 2002) offers a unique opportunity to assess the robustness of these constraints. Furthermore, the possible articulatory correlates of stable F1 values are examined within an articulatory model of the vocal tract, enabling us to link acoustic and articulatory variables in a coherent way.

2. Method

2.1. Speakers and corpus

In the first corpus, 12 native speakers of Continental French (hereafter CO) in the following age groups participated in the study: 4-year-old (two females), 8-year-old (two males and two females) and adult (three males and three females). The three groups averaged 4 years 10 months of age (3 years 10 months and 5 years 10 months), 8 years 1 month of age (from 6 years 2 months to 9 years 11 months), and 25 years of age (from 18 years to 39 years). These three groups will be referred to as the 4-year-old group, the 8-year-old group, and the adult group from the CO corpus. None of the speakers reported any history of auditory or articulatory disability. The screening procedure consisted of (1) a brief conversation with the experimenter and a speech language pathologist, (2) a 20-dB pure-tone screening at 500, 1000, 2000, 4000, and 8000 Hz, and (3) for children, a brief developmental test in order to detect speech production disabilities (*Nouvelles Études pour l'Évaluation du Langage*, Chevrie-Muller and Plaza, 2001). The corpus consisted of ten repetitions of the ten French oral vowels /i y u e ø o ɛ œ ɔ a/. Table 1 presents the feature analysis of the French vowel system. All vowels were elicited in the following forms: *V comme WORD* ('V as in WORD'), where V is one of the ten vowels mentioned above, and WORD is a French word with this vowel in initial position. Only the first vowel V, long and sustained, was analyzed. All speakers repeated the sequence after hearing an adult speaker utter it. The speech signals were recorded in a sound booth with a high-quality tabletop microphone (Sony) at a 15–20 cm distance from the subject's lips, and digitized at 44100 Hz by a Digital Audio Tape Recorder (DAT).

In the second corpus, 15 native speakers of Canadian French (hereafter CA) were recorded, from 4 years of age to adulthood. There were five speakers in each group, with

four females and one male in each of the 4-year-old and 8-year-old groups and one female and four males in the adult group. The three groups averaged 4 years 6 months of age (from 4 years to 4 years 11 months), 8 years 4 months of age (from 7 years 11 months to 9 years 1 month), and 24 years of age (from 22 years to 29 years), respectively. None of the speakers reported a history of auditory or articulatory disability. The screening and recording procedures were similar to those used for the CO corpus presented above. The same prompts were carefully pronounced by a native speaker of Canadian French (a trained phonetician).[i] As shown in Table 1, the main difference between CA and CO French in the production of isolated vowels concerns the mid-low back vowel /ɔ/. Indeed, Canadian speakers often pronounce this vowel as a low back vowel [ɒ]. This typical feature of CA French is shown in parentheses in Table 1.

2.2. Acoustic analysis

Signals were then downsampled to 22050 Hz, after low-pass filtering (cut-off frequency of 10000 Hz). Using Linear Predictive Coding (LPC) analysis, the first three formant frequencies were extracted. The number of poles varied from 10 to 14, depending on speaker age. For the 4-year-old group, since formant values are higher, and thus, fewer formants can be detected for the same sampling frequency, a coefficient of 10 was used. For the 8-year-old group and the adult group, coefficients ranging from 12 to 14 were used in the algorithm. In order to avoid automatic detection errors, all formant frequencies were overlaid on a broadband spectrogram. When a discrepancy was observed between the automatically detected values and the spectrogram, the number of

poles was adjusted and the analysis was performed again. In order to measure the validity of the data, a second experimenter randomly selected 20 tokens and manually measured F1, F2, and F3 on the spectrogram. The differences between the automatically detected formant frequencies and the manually extracted frequencies were as follows (in percentage of the mean values): 1.3% (8 Hz) for the first formant, 1.4% (29 Hz) for the second formant and 1.5% (59 Hz) for the third formant. Values were in the range of those found by Lee et al. (1999) and Hillenbrand et al. (1995). The formant frequencies were then converted to the Bark scale since this scale models the perceptual distribution of frequencies in the human auditory system, following the formula found in Schroeder et al. (1979): $F_{Bark} = 7*\text{asinh}(F_{Hz} / 650)$.

2.3. Calculation of acoustic distances between height degrees

In order to quantify the distribution of high, mid-high, mid-low, and low vowels along F1, following Neagu (1997), we compared the distance in the F1 dimension, in Bark, between vowels of different heights. High, mid-high, mid-low, and low vowels (Table I) were respectively associated with degrees 1, 2, 3, and 4. Then, the following calculations were carried out, based on the data in Bark:

- for each speaker, mean F1 values for each vowel were computed (x_j, where j is one of the ten French oral vowels /i y u e ø o ɛ œ ɔ a/);
- we defined $m_1 = (x_i+x_y+x_u)/3$ and $m_4 = x_a$ as the minimal and maximal F1 values for the speaker in question;

- a normalized index for each of the mid-high and mid-low vowels was computed by the formula: $y_j = 100*(x_j–m_1)/(m_4–m_1)$, $j \in$ /e ø o ɛ œ ɔ/.

A schematic representation is given in Figure 1. For each dialect, an ANOVA was carried out on these normalized indices (dependent variable) with height (mid-high or mid-low) and place/roundedness (front unrounded, front rounded, or back) as the within-subject fixed factors, and subject as the random factor. Interaction effects were further explored by planned comparisons using the Bonferroni correction with the alpha level set to 0.05.

2.4. Simulations with an articulatory-to-acoustic model

The next step consisted of assessing possible articulatory correlates of stable F1 values, using an articulatory-acoustic model of the vocal tract. For this study, we used the *Variable Linear Articulatory Model* (VLAM), developed by Shinji Maeda, a growth-driven scaling of an adult version of Maeda's model (Maeda, 1979), which was established on the basis of cineradiographic data and derived from a statistical analysis guided by knowledge of the physiology of the articulators. The VLAM is extensively described elsewhere (Boë, 1999; Ménard et al., 2002, to appear), and its main features will be only briefly described here. This model, controlled by seven articulatory parameters (protrusion and labial aperture; movement of the tongue body, dorsum and tip; jaw height; larynx height), generates a two-dimensional mid-sagittal section, as well as the corresponding area function (three-dimensional equivalent), from which it is possible to calculate the harmonic response (transfer function), formant frequencies (resonance maxima), and speech signal. The growth process is introduced by modifying

the longitudinal dimension of the vocal tract according to two scaling factors, one for the anterior part of the vocal tract and the other for the pharynx, interpolating the zone in between. Vocal tract shape can be simulated, month by month and year by year: this was calibrated using the data provided by Goldstein (1980). For our study, we set the model to 4 years old and 21 years old, the latter corresponding to the adult stage. These ages were chosen since they correspond to the younger and the older ages of our speakers.

The prototypical formant values of the ten French oral vowels were those used in Ménard et al. (to appear). Briefly, the ten vowels were situated within the maximal acoustic vowel space that could be generated by a combination of all possible values for the seven control parameters, using data provided by typological studies (Bailly et al., 1995; Vallée, 1994). Prototypical locations of the French oral vowels within the acoustic vowel space are depicted in Figure 1. Importantly, the locations of prototypical vowels in the (F1, F2, F3) space respects series of stable F1 positions for each height series, as Figure 1 shows. Hence, they provide a good basis for assessing how the VLAM achieves F1 stability in spite of large F2 differences between front and back vowels.

To this end, for each prototypical vowel, the articulatory parameters were inferred from a formant-to-articulatory inversion process. The method used here consists of calculating the pseudo-inverse of the Jacobian matrix (Jordan and Rumelhart, 1992). Because of the many-to-one relationship between articulatory configurations and acoustic values (e.g., Atal et al., 1978; Boë et al., 1992), an iterative procedure incorporating random search was carried out to retrieve over 50 possible articulatory configurations for

each vowel, for a given growth stage. All of these configurations produce exactly the same formants, providing a sample of different articulatory configurations compatible with the acoustic output. The basic goal here was to find possible correlates of stable F1 values in the 150 articulatory configurations corresponding to the three prototypical vowels in each height series.

3. Results

3.1. Analysis of produced vowels

Dispersion ellipses of the ten repetitions of the ten vowels, in the F1 vs. F2 spaces, are depicted in Figure 2 for Continental French and Figure 3 for Canadian French. The contours of the ellipses correspond to the probability distribution for which the covariance matrix is based on F1 and F2, at a radius of 1.5 standard deviations from the mean. For the sake of clarity, different scales were used along the F1 and F2 axes, to visually normalize between-speaker variability in age and dialect. As these figures reveal, the partition of the F1 dimension varies among the speakers. Indeed, the distribution of vowels according to height does not exploit a criterion of maximal distance. That kind of pattern, previously analyzed by Neagu (1997) with adult speakers, would be represented by an equal distance between the four height degrees, that is, high (/i y u/), mid-high (/e ø o/), mid-low (/ɛ œ ɔ/) and low (/a/) vowels. By contrast, for some speakers, the high and mid-high vowels are very close in F1 (e.g., speaker b) CO_4_f in Figure 2); for others, the mid-high and mid-low vowels are very close (e.g., speaker h) CA_8_m in

Figure 3); whereas for yet others, the mid-low and low vowels are neighbors in F1 (e.g., speaker f) CO_8_m in Figure 2). However, vowels of the same height degree tend to be realized with similar F1 values, even though the vowels vary in place of articulation and roundedness. Visually, this phenomenon is noticeable by an alignment of high (/i y u/), mid-high (/e ø o/), and mid-low vowels (/ɛ œ ɔ/) along three relatively stable F1 values.

The relative positions of the vowels along the F1 dimension (*y*-values) for each speaker of the CO corpus are presented in Figure 4. In this graph, the relative positions of mid-high and mid-low vowels along the F1 dimension are represented as a percentage of the speaker's F1 range (distance between high and low vowels). Low *y*-values (in the upper portion of the graph) corresponding to mid-high vowels reflect the small F1 distances between high and mid-high vowels, whereas high *y*-values for mid-low vowels stand for the large F1 distances between high and mid-low vowels. For each speaker, the values for the three mid-high vowels are linked by a solid line, whereas the values for the three mid-low vowels are linked by a dotted line. Long solid or dotted lines denote large within-speaker variation in the F1 values within a given height degree. Speakers are sorted along the *x*-axis in ascending order of their *y*-data points for mid-high vowels.

It is striking to observe, first, the great between-speaker variability in the relative position of mid-high vowels, as depicted by the position of the solid lines along the *y*-axis. Indeed, values range from 2 (minimal value) to 47 (maximal value), resulting in a between-speaker variability of 45. The same between-speaker variability pattern is observed for mid-low vowels. *Y*-values for these data points range from a minimal value

of 27 to a maximal value of 91 (for a between-speaker variability of 64). No effect of age is found, as revealed by the fact that speakers from all three age groups are associated with both small and large *y*-values. Furthermore, *y*-values for mid-high (/e ø o/) and mid-low vowels (/ɛ œ ɔ/) do not follow the prediction of the maximal acoustic distance criterion. Indeed, these values do not correspond to 33 and 67, which would mean that vowels of different height degrees are equally spaced along the F1 dimension.

This seemingly very variable organization among speakers, however, shows great coherence within the place of articulation and roundedness features. Recall that the length of the solid and dotted lines in Figure 4 represents the within-speaker variation in F1 for vowels of the same height. It is striking to observe that for all speakers, the *y*-values of the three mid-high vowels (length of the solid line) fall within a range of 10, a value much smaller than the between-speaker variability in the relative F1 values reported above (45). A similar pattern is found for the mid-low vowels. Indeed, for ten speakers, the within-speaker range of relative F1 values for the mid-low vowel series is below 15, and it is lower than 28 for all speakers; again, this is smaller than the between-speaker variability reported above (64). A mixed ANOVA conducted on the *y*-values with height and place/roundedness as within-subject fixed factors and subject as a random factor revealed a significant effect of height ($F(1,11) = 127.18$; $p < .05$), with mid-high vowels having higher values than mid-low vowels, as expected. No effect of place/roundedness was found, as a main effect or in interaction with height. A significant effect of the subject factor was found ($F(11,11.6) = 3.97$; $p < .05$), suggesting that F1 distances between height degrees are speaker-specific.

A similar pattern is found for the mid-high vowels of the 15 speakers from the CA corpus. These data are depicted in Figure 5. An examination of the relative position of mid-high vowels, as depicted by the position of the solid lines along the *y*-axis, reveals significant between-speaker variability. Indeed, values range from 2 to 39, for a between-speaker variability of 37. As regards within-speaker variability, for 12 speakers out of the 15, the *y*-values of the three mid-high vowels are within a range of 10, and the range is less than 18 for all speakers, which is lower than the between-speaker variability (37). For mid-low vowels, minimal and maximal *y*-values for these data points range from 18 to 58. As was the case for the CO corpus, no effect of age is found, as revealed by the fact that speakers from all three age groups are associated with both small and large *y*-values. However, within-speaker variability of mid-low vowels (dotted lines) shows a somewhat different pattern, with some speakers producing one mid-low vowel at a greater distance from the high vowels compared to the other two mid-low vowels. Detailed examinations of the values reveal that this pattern can be ascribed to the realization of the mid-low vowel /ɔ/ as the low vowel [ɒ], which is typical of Canadian French, as mentioned earlier. If /ɔ/ is discarded, once again the within-speaker variability along the *y*-axis for a given height degree (length of the solid and dotted lines) is smaller than the between-speaker variability.

The results of a mixed ANOVA carried out on all the *y*-values depicted in Figure 5 (including /ɔ/) with height and place/roundedness as fixed within-subject factors and subject as a random factor show a significant effect of height ($F(1,14) = 94.27$, $p < .05$),

with mid-low vowels having higher values than mid-high vowels, as expected. Place/roundedness also has a significant effect on *y*-values ($F(2,28) = 8.67$; $p < .05$). The interaction of height and place/roundedness is significant ($F(2,28) = 19.02$, $p < .05$). This effect can be ascribed to the phonological behavior of the mid-low back vowel /ɔ/, realized as a low back vowel [ɒ]. A significant effect of the interaction between height and subject is found ($F(14,28) = 3.66$; $p < .05$), revealing the speaker-specificity of F1 values. More importantly, no effect of the interaction between place of articulation and subject, or between height, place of articulation and subject, is found. When the same ANOVA is performed without /ɔ/, a significant effect of height is observed as a main effect ($F(1,14) = 44.44$; $p < .05$) and in interaction with the subject factor ($F(14,14) = 8.71$; $p < .05$). The effect of place/roundedness is not significant, as a main effect or in interaction with the height factor.

The latter finding clearly confirms that the distances between height degrees *y* are variable across speakers but, apart from /ɔ/, speakers tend to align all vowels of a given height on the same F1 axis. Inter-subject variations appear differently in the two corpora, with a subject effect in one case and a subject*height interaction effect in the other. The difference is basically due to the fact that, in the CO corpus (Figure 4), the variations between F1 values for the mid-high and mid-low vowels seem to be more correlated than in the CA corpus (Figure 5). However, the effect is globally similar, with large variations from one speaker to another, small differences in F1 for a given height series, and various patterns of close high and mid-high, mid-high and mid-low, or mid-low and low series in both corpora. These results are very similar to Neagu's (1997) findings, based on male

and female adult speakers. The potential articulatory or geometrical nature of these targets will be discussed in the next section.

3.2. Keeping stable F1 values in VLAM

The 50 articulatory configurations corresponding to each vowel were analyzed in order to determine whether or not the stable F1 values are related to stable underlying articulatory gestures or geometrical configurations. Recall that these modeled vowels are based on an alignment of vowels with similar height degrees along stable F1 values. Even though the specific F1 distance between height degrees follows from the criterion of maximal distance, this pattern represents our speakers' data quite well. The target F1 values for each height degree are as follows, for the 4-year-old vocal tract: 440 Hz (4.4 Bark), 630 Hz (6 Bark), and 850 Hz (7.6 Bark), respectively, for the high, mid-high, and mid-low vowels[1]. Target F1 values for the corresponding height degrees modeled for the 21-year-old vocal tract (adult) are 245 Hz (2.6 Bark), 365 Hz (3.7 Bark), and 495 Hz (4.9 Bark). Note that those values differ from those produced by our speakers (Figures 2 and 3), since the model simulates an average vocal tract length for a given growth stage.

Theoretically, the articulatory parameters involved in variation along the F1 dimension and mainly related to the openness feature are jaw height, tongue body and tongue dorsum positions. In turn, these parameters contribute to the value of the

[1] It has to be noted that those values may differ from those measured in Figures 2 and 3 due to variations in vocal tract length. The target values in the model are chosen according to the synthesized vowel space for a given speaker (Ménard *et al.*, 2004).

constriction area (geometrical parameter), and of the tongue's highest position, both of which are closely linked with F1 (Boë et al., 1992). The values of the jaw parameter, for high vowels (/i y u/), mid-high vowels (/e ø o/), and mid-low vowels (/ɛ œ ɔ/) are plotted in Figure 6 for a 4-year-old and an adult vocal tract. Basically, the distributions of values of this articulatory parameter are both very broad and quite overlapping. This is not surprising, considering the very wide range of possible compensation mechanisms involved in height control (see, for example, the bite block experiments in Lindblom et al., 1979). Values for the constriction areas are also depicted in Figure 6. The dispersion of constriction values is much reduced in comparison to the jaw parameter. Once again, this fits quite well with both experimental data on compensatory articulation (Gay et al., 1981), and articulatory-acoustic modeling showing the crucial role of constrictions in the control of acoustics (Boë et al., 1992). However, the distribution of constriction area values does not constitute a very systematic link with stable F1 values, mainly because of back vowels, which tend to have overlapping constriction area ranges, particularly /u/ and /ɔ/ for the 4-year-old model and /ɔ/ and /o/ for the adult one. A discriminant analysis carried out on the constriction area values with height (high, mid-high, and mid-low) as the grouping factor revealed that this geometrical parameter achieves mean percentages of correct classification of 82% and 83% for the 21-year-old and 4-year-old vocal tracts, respectively.

In order to better visualize the articulatory strategies involved in the synthesized vowels in both simulated vocal tracts, a schematic representation of tongue positions is provided in Figure 7. Each dispersion ellipsis corresponds to the *xy* coordinates of the

highest point of the tongue for a given vowel category. The ellipses are drawn at a radius of ±1.5 standard deviations from the mean. For the sake of clarity, the VLAM palate traces are superimposed on the graphs. Figure 7 shows that height degrees are not associated with specific *x* or *y* coordinates, but correspond to distinct regions within that space. Specifically, the position of the front and back vowels in the space seems to parallel the palate surface, more or less, as indicated by the dotted lines drawn on the graphs. The values for the 50 articulatory configurations are rather tightly clustered within each vowel category, as was the case for constriction areas, but they are better separated than the latter. Since it has been shown that similar height degrees tend to be produced by stable F1 values, the distribution of height degrees in the *xy* space, according to the lines superimposed on the figure, may provide an acceptable correlate of F1. Hence, the distance between the highest point of the tongue and the palate provides a possible characterization of stable F1 values. In order to quantify the extent to which such a space allows different heights to be adequately distinguished, a discriminant analysis was carried out with *x* and *y* coordinates as the classification parameters and height (high, mid-high, and mid-low) as the grouping factor. The average percentage of correct classification scores reached 89% and 80% for the 21-year-old and 4-year-old vocal tracts, respectively; these values are in the range of those found with constriction area as the classification parameter (respectively, 82% and 83% for the 21-year-old and 4-year-old vocal tracts). As mentioned above, considering the fact that the distributions follow the surface of the palate (which is ogive-shaped), a stable tongue distance relative to the palate is represented here.

4. Discussion

The results of this experiment display both invariance and variability in the distribution of vowels within the two variants of the French system that we studied. Indeed, in our 4-year-old, 8-year-old, and adult subjects, the partition of the F1 dimension is very stable across place of articulation and roundedness, within each speaker's system. It is important to note that this stable pattern was observed for all speaker groups (regardless of age and dialect, apart from the specific case of /ɔ/), which reveals its robustness. F1 values associated with height are speaker-specific and do not conform to a maximal contrast criterion between height degrees. However, the results can be interpreted in light of the PACT theory, in which speech units are considered as speech gestures shaped by perceptual processes.

4.1. Stable F1 values: why and how?

The rather strong stability of F1 values in height series, in spite of their obvious variability from speaker to speaker, calls for an explanation. First of all, it cannot be given a direct phonological interpretation, because of the inter-speaker variability. For the same reason, this behavior cannot be conceived of as being learned by one speaker from another in the course of development. It could be argued that young speakers observe the constancy of F1 in spite of its variations, and then learn to reproduce this behavior, but this still would not tell us why speakers of a given language community obey this constancy law. Furthermore, no theory of speech communication relies on these kinds of

metaphonetic ingredients, such as “two or more phonemes share the same spectro-temporal characteristic, whatever its value is.”

There is no obvious listener-oriented gain in F1 stability, considering once again that inter-subject variability would prevent such stability from being exploited for phoneme or feature identification, all the more so considering the dramatic changes in formant values displayed in normal speech utterances because of reduction and coarticulation.

Therefore, the reason must be speaker-oriented. The most plausible assumption is that keeping F1 values stable in a given series simplifies vowel control, and probably vowel learning in development. At this level, two hypotheses might be proposed concerning the potential benefits for control. The first is that the stability of F1 values per se may play a role in mastering the French vowel system. This may be compatible with auditory theories of speech production, whereby speech targets are directly specified in auditory terms (Perkell et al., 1997, 2004), possibly through articulatory-auditory maps and learned inversion mechanisms (Bailly, 1997; Guenther, 1995; Guenther et al., 1998). In this framework, it is not impossible to imagine that such articulatory-auditory maps would be simplified if F1 values remained stable, thus easing, at some level, the speech production process.

More likely in our view is a second assumption, namely that a more proximal sensorimotor variable is stabilized by the constancy of the F1 series. In this context,

tongue height appears to be an appealing candidate, considering the reasonably good correlation between F1 values and tongue height or tongue-palate proximity discussed in section 3.2. In a set of perceptual experiments using stimuli from VLAM, Vallée and Kandel (2003) tested the ability of naive French listeners to determine which of a given pair of stimuli corresponded to the higher vs. lower tongue height configuration. They showed that even naive listeners seemed quite able to deduce vowel aperture from speech sounds, whatever the front-back tongue position or labial configuration, and even when the vowel did not belong to the French phonological system. Furthermore, F1 was the basic correlate of the subjects' performance. This suggests that speakers are reasonably aware of the correspondence between F1 and tongue height and therefore are able to maintain stable F1 values in order to achieve a series of tongue configurations that they feel to be stable in terms of palatal proximity. This may simplify the somatosensory feedback needed to control the speech task (Ostry and Nasir, 2006; Tremblay et al., 2003). This hypothesis could be further investigated using speech synthesis (to control F1 and F2 variation) and perceptual experiments.

4.2. Interpretation of the results within the PACT framework

The results presented in this paper suggest that the position of vowels along the F1 dimension (height degree) is not shaped by a constraint aimed at maximizing the acoustic distance between vowels. Indeed, the relative distance in F1 between high and mid-high vowels and between high and mid-low vowels seldom corresponds to 33 or 67 (Figures 4 and 5). Thus, contrasting vowels along the height dimension are not equally spaced along

the F1 acoustic parameter. The acoustic contrast is nevertheless sufficient, according to Lindblom (Lindblom, 1986; Lindblom and Maddieson, 1988).

The fact that acoustic dispersion need not be maximal in order to be sufficient, as shown by our data, does not explain why vowels of a similar height tend to be aligned along the same F1 value, within the vowel space. The same limitation was found by Schwartz et al. (1997), who showed that the dispersion constraint (global) and the focalization constraint (local) could not account for the fact that vowel systems in the world's languages tend to equilibrate peripheral vowels: if a system features a mid-high front vowel /e/, then it also features the mid-high back counterpart /o/, not the mid-low back vowel /ɔ/. Similarly, if a vowel system contains the peripheral vowel /ɛ/, a mid-low vowel, then it generally contains the mid-low counterpart /ɔ/. Obviously, perceptual dispersion accounts of the acoustic organization of the vowel space cannot explain such regularities. Rather, we suggest that this structural pattern can be accounted for within the PACT (Schwartz et al., 2006). Recall that this theory assumes that vowel systems are organized following dispersion constraints, regularized by articulatory knowledge. The regularity principle here operates at the level of articulatory controls, rather than features, as in the MUAF (Ohala, 1979). This hypothesis is coherent with current theories of speech ontogeny.

In the course of speech acquisition, according to the Frame Then Content theory (MacNeilage and Davis, 1990), the oscillatory movement of the jaw is the first degree of freedom controlled by the baby. The open-close alternation cycles of the jaw give rise to

consonant-like sounds in the upper part of the cycle and vowel-like sounds in the lower part. This control provides the framework for the subsequent emergence of segmental control and differentiation. The first set of sounds produced by the baby at 6 to 8 months old mostly belong to the front region of the vowel space: mid-high, mid-low, and low vowels. The control of tongue height, which allows contrasts along the height feature, is progressively acquired. In a first approximation, it can be proposed that when control over tongue and lip movements is acquired, new contrasts can be produced. Since specific tongue heights and the corresponding sensorimotor controls are already associated with front vowels, it would be more economical and easier to learn to combine the new articulatory controls with those already available, namely specific tongue height positions. Thus, the progressive tuning of specific height controls would occur globally, for all places of articulation and lip rounding configurations. This would progressively lead to a restricted set of tongue-palate distances, selected by the speaker as being both articulatorily adequate (in terms of stable somatosensory feedback) and perceptually sufficiently contrasted. The tendency to align vowels of similar heights along configurations with more or less stable tongue-palate proximity (resulting in a stable F1 value) would thus be related to economy of articulatory control.

5. Conclusion

This paper investigated the partition of the F1 dimension along different height degrees in French oral vowels. Acoustic recordings of French oral vowels produced by children and adults who spoke two dialects were analyzed. The data showed that although the specific

F1 values corresponding to a given height degree are speaker-dependent, speakers tend to align vowels of a similar height along stable F1 values across rounding and place of articulation. Comparisons with simulations using an articulatory-acoustic model revealed that maintaining stable F1 patterns for similar height degrees apparently involves adaptive articulatory gestures for front rounded and unrounded vowels and back vowels. The organization of the vowel space thus shows both invariance, in the sense of great stability for a given height value and a given speaker, and variability, with large F1 differences from one speaker to another. This pattern cannot be accounted for by previous theories of vowel systems based on perceptual constraints alone, and also departs from a strictly formal MUAF principle. Rather, these data are in line with the PACT, in which the sound systems of human languages are considered as the result of a perceptuo-motor link in which speech units are produced by perceptual distinctiveness and focalization constraints, but regularized by articulatory control principles allowing the speaker to master the communication system conveniently and idiosyncratically.

---

[i] Even though we cannot rule out the possibility of slight variations in the pronunciation of the prompts by the experimenters from one recording session to the other, we believe those variations, if any, did not influence the pattern of results presented here.

Acknowledgments

This work was supported by grants from the Social Sciences and Humanities Research Council of Canada, the Natural Sciences and Engineering Research Council of Canada, and the Fonds Québécois de la Recherche sur la Société et la Culture. We are indebted to Shinji Maeda, Willy Serniclaes, and Marija Tabain for fruitful discussions and helpful suggestions. The authors thank Zofia Laubitz for copy-editing the article.

REFERENCES

Atal, B.S., Chang, J.J., Mathews, M.V., Tukey, J.W., 1978. Inversion of articulatory-to-acoustic transformation in the vocal tract by a computer-sorting technique. J. Acoust. Soc. Am. 63, 1535-1555.

Bailly, G., 1997. Learning to speak. Sensori-motor control of speech movements. Speech Commun. 22, 251-267.

Bailly, G., Boë, L.-J., Vallée, N., Badin, P., 1995. Articulatory-acoustic vowel prototypes for speech production. Proc. 4th Eur. Conf. Speech Commun. and Tech., Madrid, 3, 1913-1916.

Boë, L.-J., 1999. Modelling the growth of the vocal tract vowel spaces of newly-born infants and adults. Consequences for ontogenesis and phylogenesis. Proc. Intern. Congress Phon. Sci., 3, San Francisco, 2501-2504.

Boë, L.-J., Perrier, P., Bailly, G., 1992. The geometric vocal tract variables controlled for vowel production: Proposals for constraining acoustic-to-articulatory inversion. J. Phonetics, 20, 27-38.

Chevrie-Muller, C., Plaza, M., 2001. Nouvelles Études pour l'Examen du Langage. Ecole de Psychologie Appliquée, Paris.

Gay, T., Lindblom, B., Lubker, J., 1981. Production of bite-block vowels: Acoustic equivalence by selective compensation. J. Acoust. Soc. Am. 69, 802-810.

Goldstein, U.G., 1980. An articulatory model for the vocal tract of the growing children. Doctoral dissertation, MIT, Cambridge, MA.

Guenther, F.H., 1995. Speech sound acquisition, coarticulation, and rate effects in a neural network model of speech production. Psychol. Rev. 102, 594-621.

Guenther, F.H., Hampson, M., Johnson, D., 1998. A theoretical investigation of reference frames for the planning of speech movements. Psychol. Rev. 105, 611-633.

Hillengrand, J., Getty, L. A., Clark, M. J., Wheeler, K. 1995. Acoustic characteristics of American English vowels. J. Acoust. Soc. Am. 97, 3099-3111.

Jordan, M.I., Rumelhart, D.E., 1992. Forward models: Supervised learning with a distal teacher. Cogn. Sci. 16, 316-354.

Lee, S., Potamianos, A., Narayanan, S., 1999. Acoustics of children's speech: Developmental changes of temporal and spectral parameters. J. Acoust. Soc. Am. 105, 1455-1468.

Liberman, A.M., Mattingly, I.G., 1985. The motor theory of speech production revised. Cognition. 21, 1-36.

Lindblom, B., 1986. Phonetic universals in vowel systems, in: Ohala, J.J., Jaeger, J.J. (Eds.), Experimental Phonology, Academic Press, New York, pp. 13-44.

Lindblom, B., 1996. Role of articulation in speech perception: Clues from production. J. Acoust. Soc. Am. 99, 1683-1692.

Lindblom, B., 1998. Systematic constraints and adaptive change in the formation of sound structure, in: Hurford, J.R., Studdert-Kennedy, M., Knight, C. (Eds.), Approaches to the Evolution of Language, Cambridge University Press, Cambridge, pp. 242–264.

Lindblom, B., Lubker, J., Gay, T., 1979. Formant frequencies of some fixed mandible vowels and a model of speech motor programming by predictive simulation. J. Phon. 7, 147-161.

Lindblom, B., Maddieson, I.,1988. Phonetic universals in consonant systems, in: Hyman, L.H., Li, C.N. (Eds.), Language, Speech, and Mind, Routledge, London, pp. 62-78.

MacNeilage, P.F., Davis, B. L., 1990. Acquisition of speech production: Frames then content, in: Jannerod, M. (Ed.), Attention and Performance XIII: Motor Representation and Control, Lawrence Erlbaum, Hillsdale, NJ, pp. 453-475.

Maeda, S., 1979. An articulatory model of the tongue based on a statistical analysis. J. Acoust. Soc. Am. 65, S22.

Ménard, L., 2002. Production et perception des voyelles au cours de la croissance du conduit vocal: variabilité, invariance et normalisation. Doctoral dissertation, Grenoble, Université Stendhal/Institut de la communication parlée.

Ménard, L., Schwartz, J.-L., Boë, L.-J., 2004. Role of vocal tract morphology in speech development: Perceptual targets and sensori-motor maps for synthesized French vowels from birth to adulthood. J. Speech Lang. Hear. Res. 47, 1059-1080.

Ménard, L., Schwartz, J.-L., Boë, L.-J., Aubin, J., to appear. Production-perception relationships during vocal tract growth for French vowels: Analysis of real data and simulations with an articulatory model. J. Phon.

Ménard, L., Schwartz, J.-L., Boë, L.-J., Kandel, S., Vallée, N., 2002. Auditory normalization of French vowels synthesized by an articulatory model simulating growth from birth to adulthood. J. Acoust. Soc. Am. 111, 1892-1905.

Neagu, A., 1997. Analyse articulatoire du signal de parole: caractérisation des syllabes occlusive-voyelle en français. Doctoral dissertation, Signal-Image-Word specialization, Grenoble, INPG.

Nearey, T.M., 1997. Speech perception as pattern recognition. J. Acoust. Soc. Am. 101, 3241–3254.

Ohala, J.J., 1979. Moderator's introduction to symposium on phonetic universals in phonological systems and their explanation. Proc. Intern. Congress Phon. Sci., 3, pp. 181–185.

Ostry, D.J., Nasir, S., 2006. The somatosensory precision requirements of speech production. Proceedings of the 5th International Conference on Speech Motor Control, Nijmegen, 14.

Perkell, J., Guenther, F., Lane, H., Matthies, M., Stockmann, E., Tiede, M., Zandipour, M., 2004. The distinctness of speakers' productions of vowel contrasts is related to their discrimination of the contrasts. J. Acoust. Soc. Am. 116, 2338-2344.

Perkell, J.S., Matthies, M.L., Lane, H., Guenther, F.H., Wilhelms-Tricarico, R., Wozniak, J., Guiod, P., 1997. Speech motor control: Acoustic goals, saturation effects, auditory feedback and internal models. Speech Commun. 22, 227-250.

Perrier, P., 2005. Control and representations in speech production. ZASPIL - ZAS Papers in Linguistics. Special issue on speech production and perception: Experimental analyses and models (Fuchs, S., Perrier P., Pompino-Marschall, B., Eds.), 40, pp. 190-132.

Polka, L., Bohn, O.-S., 2003. Asymmetries in vowel perception. Speech Commun. 41, 221–231.

Schroeder, M. R., Atal, B. S., Hall, J. L., 1979. Objective measure of certain speech signal degradations based on masking properties of human auditory perception, in: Lindblom, B., Öhman, S. (Eds.), Frontiers of Speech Communication Research, Academic Press, London, pp. 217-229.

Schwartz, J.-L., Abry, C., Boë, L.-J., Cathiard, M.A., 2002. Phonology in a theory of perception-for-action-control, in: Durand, J., Laks, B. (Eds.), Phonetics, Phonology and Cognition, Oxford University Press, Oxford, pp. 255-281.

Schwartz, J.L., Abry, C., Boë, L.J., Ménard, L., Vallée, N., 2005. Asymmetries in vowel perception, in the context of the Dispersion-Focalisation Theory. Speech Commun. 45, 425–434.

Schwartz, J.-L., Boë, L.-J., Abry, C., 2006. Linking the Dispersion-Focalization Theory (DFT) and the Maximum Utilization of the Available Distinctive Features

(MUAF) principle in a Perception-for-Action-Control Theory (PACT), in: Solé, M.J., Beddor, P., Ohala, M. (Eds.), Experimental Approaches to Phonology, Oxford University Press, Oxford (to appear).

Schwartz, J.-L., Boë, L.-J., Vallée, N., Abry, C., 1997. The Dispersion-Focalization Theory of vowel systems. J. Phon. 25, 255-286.

Stevens, K. N., 1989. On the quantal nature of speech. J. Phon. 17, 3-45.

Tremblay, S., Shiller, D.M., Ostry, D.J., 2003. Somatosensory basis of speech production. Nature. 423, 866-869.

Vallée, N., 1994. Systèmes vocaliques: de la typologie aux prédictions. Doctoral dissertation, Grenoble, Université Stendhal/Institut de la communication parlée.

Vallée, N., Kandel, S., 2003. Can we recover vowel gestures from speech sounds? An experimental study based on an original psychophysical paradigm. Proc. Intern. Congress Phon. Sci., Barcelona, 817-820.

FIGURE CAPTIONS:

Figure 1: Schematic representation of the metric used to evaluate between-speaker and within-speaker variability in F1 distances between degrees of vowel height. Prototypical locations of the French oral vowels in the vowel space are shown.

Figure 2: Dispersion ellipses (±1.5 standard deviations from the mean) of the ten French oral vowels /i y u e ø o ɛ œ ɔ a/ in the F1 vs. F2 space, for the speakers of the CO corpus. Each speaker is labeled by a code of the form DIALECT_GROUP_GENDER (real age).

Figure 3: Dispersion ellipses (±1.5 standard deviations around the mean) of the ten French oral vowels /i y u e ø o ɛ œ ɔ a/ in the F1 vs. F2 space, for the speakers of the CA corpus. Each speaker is labeled by a code of the form DIALECT_GROUP_GENDER (real age).

Figure 4: Mean values of relative position along F1 (as a % of the F1 difference between high vowels and /a/) for the 12 speakers of the CO corpus. Data are presented separately for mid-high (/e ø o/, solid line) and mid-low vowels (/ɛ œ ɔ/, dotted line). y_j is calculated as $(x_j–m_1)/(m_4–m_1)*100$, where $m_1 = (x_i+x_y+x_u)/3$, $m_4 = x_a$ and j is one of the six French oral vowels /e ø o ɛ œ ɔ /, for each speaker. For a given height degree and a given speaker, the y-values of the three vowels are linked by a vertical bar. Speakers are sorted along the x-axis in ascending order of their y-data points for mid-high vowels.

Figure 5: Mean values of relative position along F1 (as a % of the F1 difference between high vowels and /a/) for the 15 speakers of the CA corpus. Data are presented separately for mid-high (/e ø o/, solid line) and mid-low vowels (/ɛ œ ɔ/, dotted line). y_j is calculated as $(x_j–m_1)/(m_4–m_1)*100$, where $m_1 = (x_i+x_y+x_u)/3$, $m_4 = x_a$ and j is one of the six French oral vowels /e ø o ɛ œ ɔ /, for each speaker. For a given height degree and a given speaker, the *y*-values of the three vowels are linked by a vertical bar. Speakers are sorted along the *x*-axis in ascending order of their *y*-data points for mid-high vowels.

Figure 6: Mean values and standard deviations of constriction area (left) and jaw position (right) for 50 articulatory configurations per French vowel simulated in VLAM for the 4-year-old vocal tract (top) and the adult male vocal tract (bottom). The high vowels /i y u/ are represented by the circles and the solid black line, the mid-high vowels /e ø o/ correspond to the triangles and the dashed blue line, and the mid-low vowels /ɛ œ ɔ/ are depicted by the squares and the solid red line.

Figure 7: Coordinates of the highest point of the tongue for 50 articulatory configurations per French vowel simulated in VLAM for the 4-year-old vocal tract (left) and the adult male vocal tract (right). Palate traces are superimposed on the graphs. H = high vowels; MH = mid-high vowels; ML = mid-low vowels.

TABLE I: Feature analysis of French vowels. Characteristics of CA French are shown in parentheses. See text for details.

	Front		**Back**
	Unrounded	**Rounded**	
High	i	y	u
Mid-high	e	ø	o
Mid-low	ɛ	œ	ɔ (ɒ)
Low	a		

FIGURE 1:

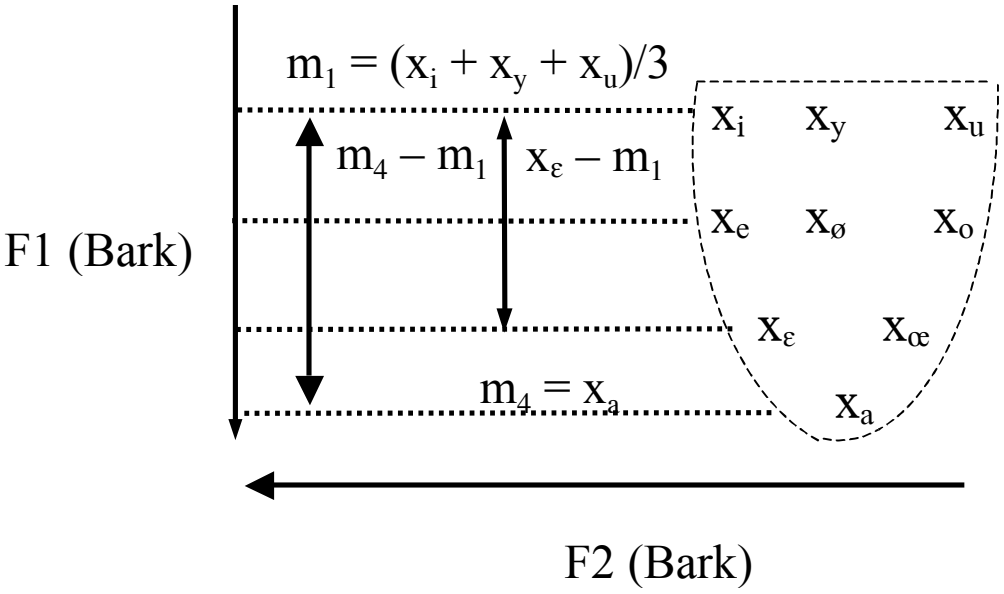
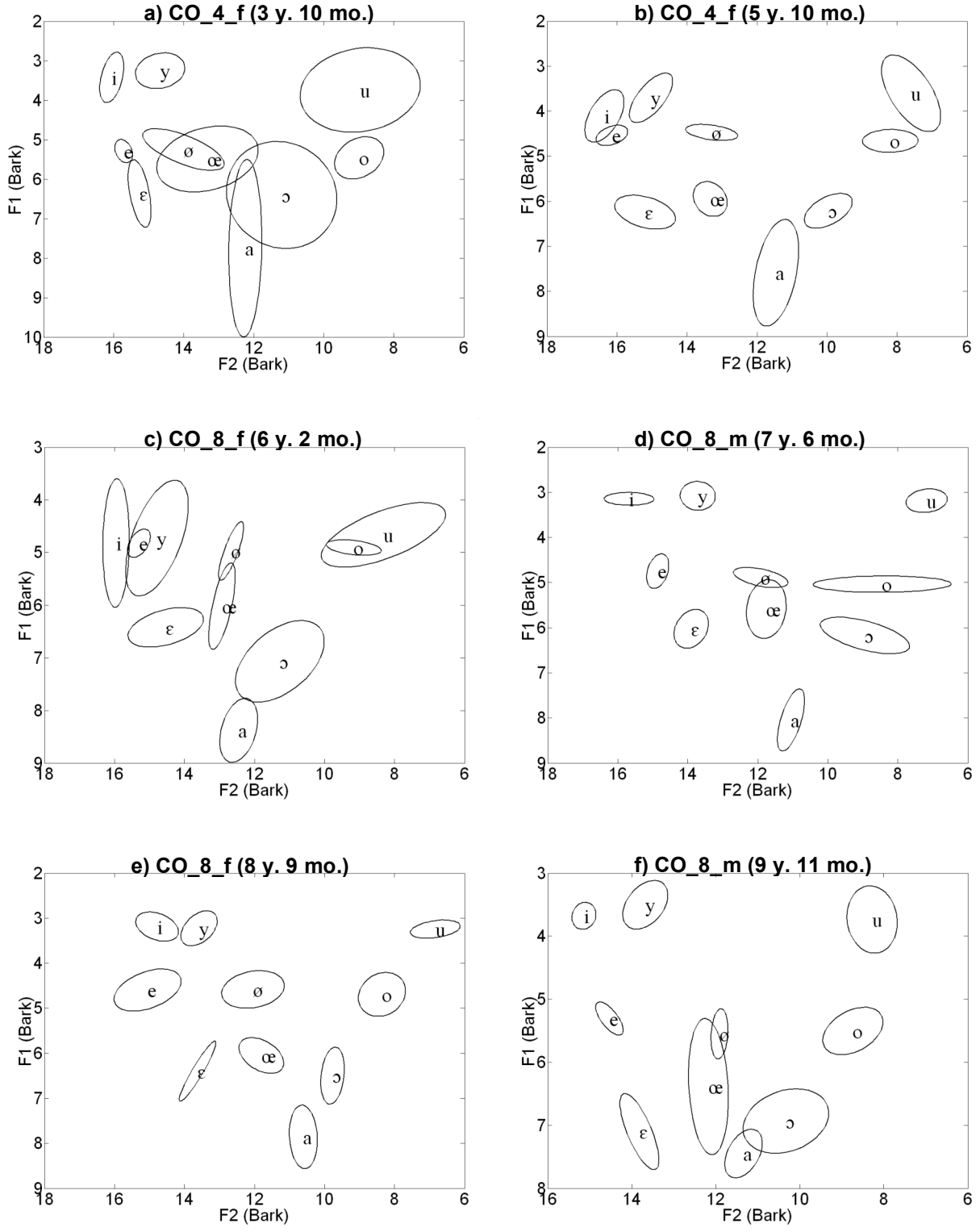
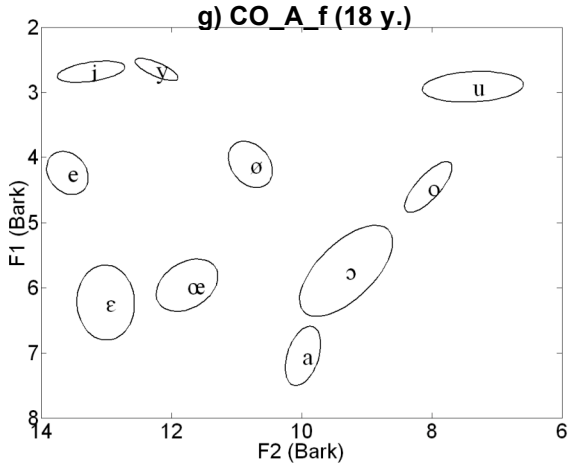
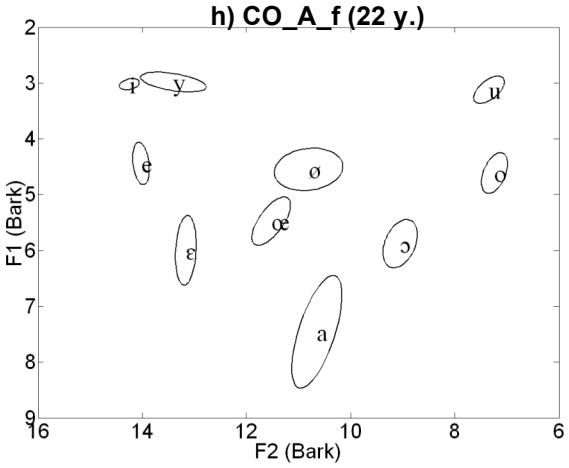
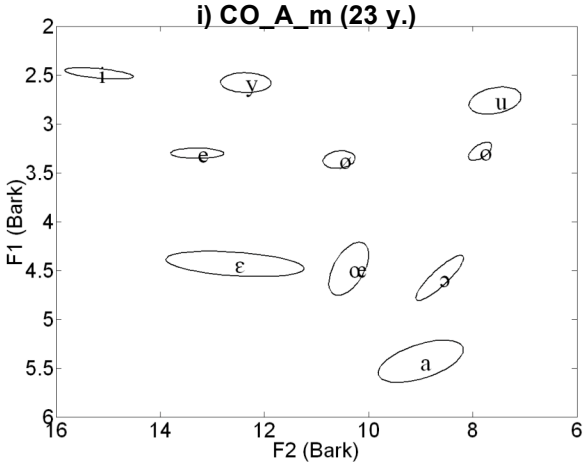

$m_1 = (x_i + x_y + x_u)/3$
$m_4 - m_1$
$x_\varepsilon - m_1$
x_i
x_y
x_u
x_e
$x_ø$
x_o
x_ε
$x_{œ}$
$m_4 = x_a$
x_a
F1 (Bark)
F2 (Bark)

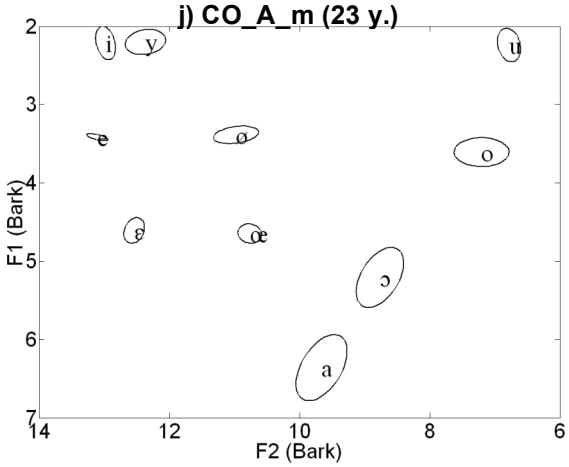
FIGURE 2:

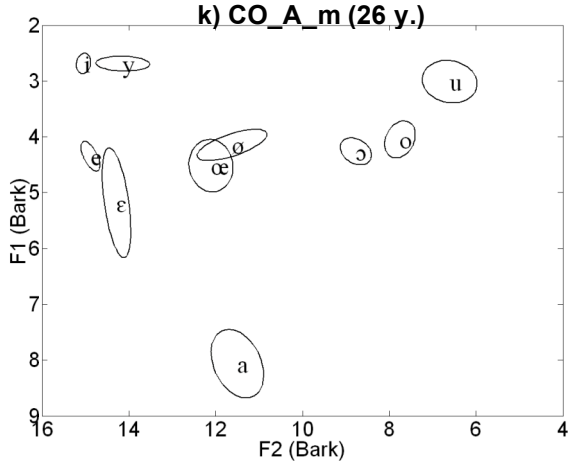

a) CO_4_f (3 y. 10 mo.)
b) CO_4_f (5 y. 10 mo.)
c) CO_8_f (6 y. 2 mo.)
d) CO_8_m (7 y. 6 mo.)
e) CO_8_f (8 y. 9 mo.)
f) CO_8_m (9 y. 11 mo.)
F1 (Bark)
F2 (Bark)
i
y
u
e
ø
œ
o
ɛ
ɔ
a


g) CO_A_f (18 y.)
F1 (Bark)
F2 (Bark)
i
y
u
e
ø
o
ɛ
œ
ɔ
a


h) CO_A_f (22 y.)
F1 (Bark)
F2 (Bark)
y
u
e
ø
o
œ
ɛ
ɔ
a


i) CO_A_m (23 y.)
F1 (Bark)
F2 (Bark)
i
y
u
e
ø
ɛ
œ
ɔ
a


j) CO_A_m (23 y.)
F1 (Bark)
F2 (Bark)
i
y
u
e
ø
o
ɛ
œ
ɔ
a


k) CO_A_m (26 y.)
F1 (Bark)
F2 (Bark)
i
y
u
e
ø
ɔ
o
œ
ɛ
a

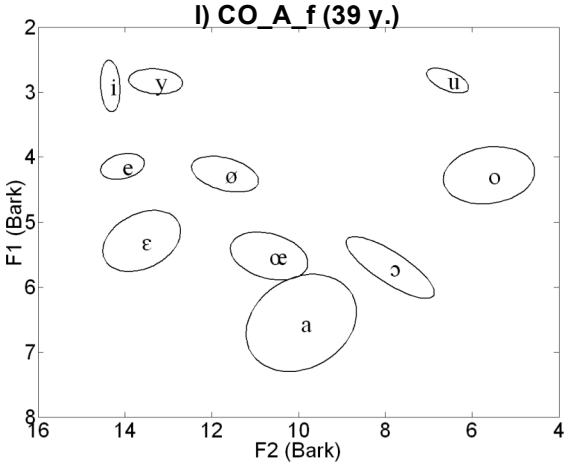
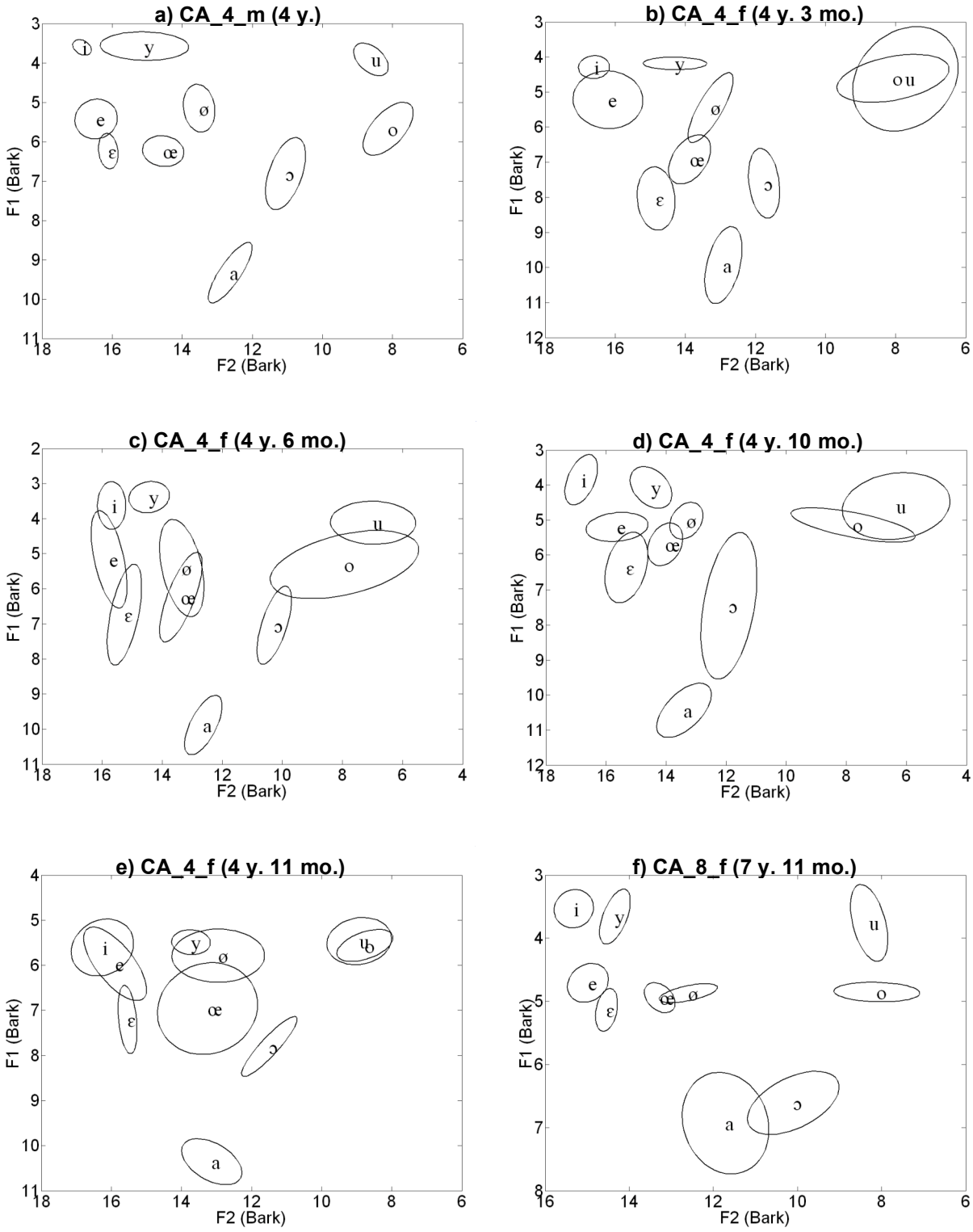
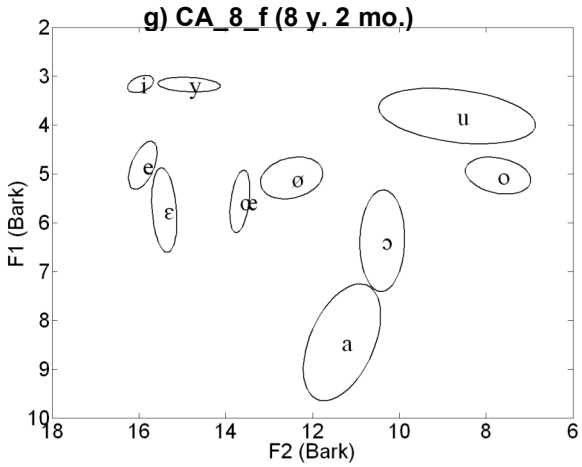
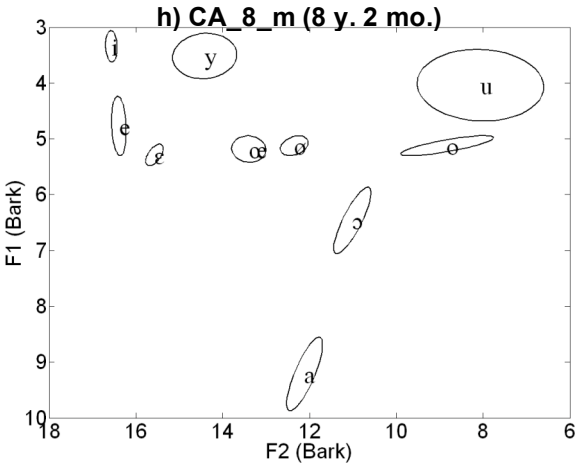
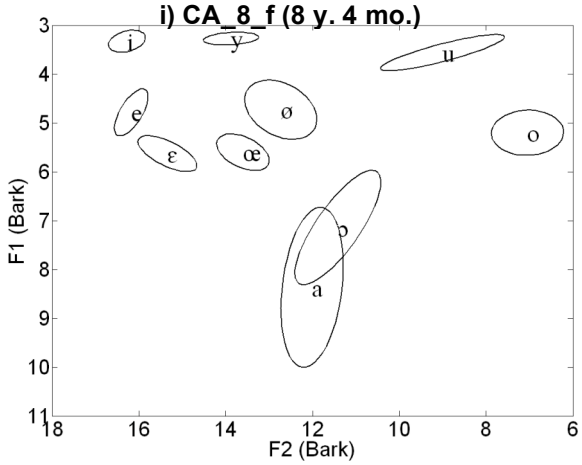

l) CO_A_f (39 y.)
F1 (Bark)
F2 (Bark)
i
y
u
e
ø
o
ɛ
œ
ɔ
a

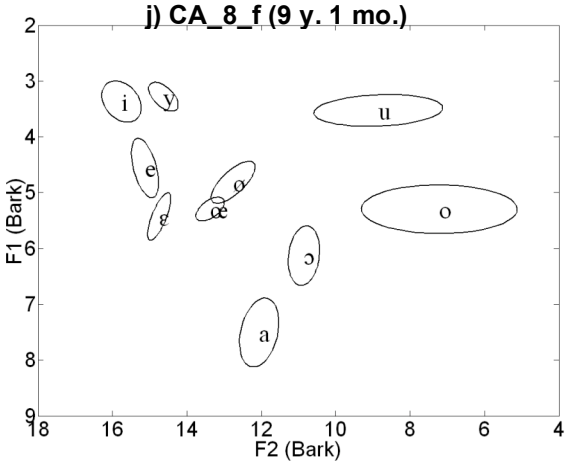
FIGURE 3:

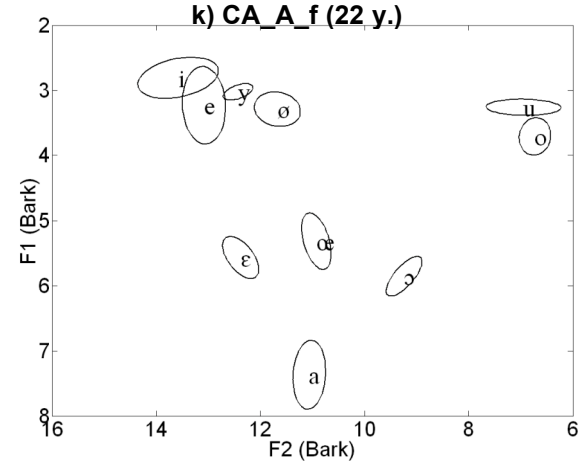

a) CA_4_m (4 y.)
b) CA_4_f (4 y. 3 mo.)
c) CA_4_f (4 y. 6 mo.)
d) CA_4_f (4 y. 10 mo.)
e) CA_4_f (4 y. 11 mo.)
f) CA_8_f (7 y. 11 mo.)
F1 (Bark)
F2 (Bark)

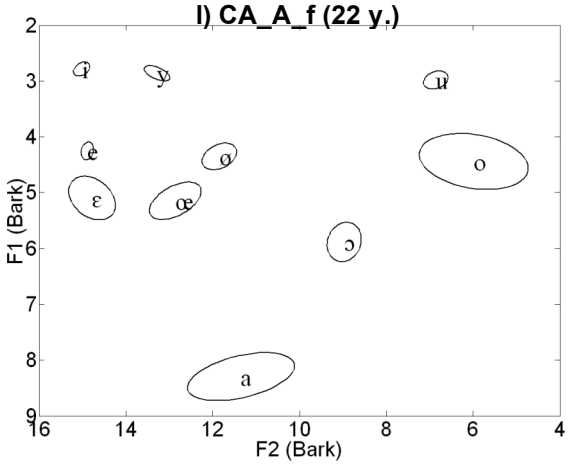

g) CA_8_f (8 y. 2 mo.)
F1 (Bark)
F2 (Bark)
i
y
u
e
ø
ɛ
œ
o
ɔ
a

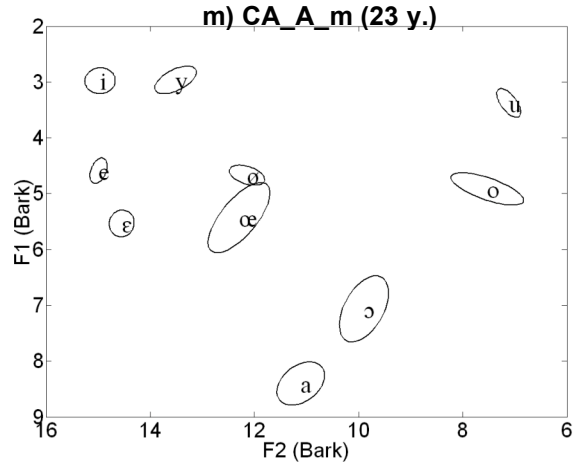

h) CA_8_m (8 y. 2 mo.)
F1 (Bark)
F2 (Bark)
i
y
u
e
ɛ
œ
ø
o
ɔ
a

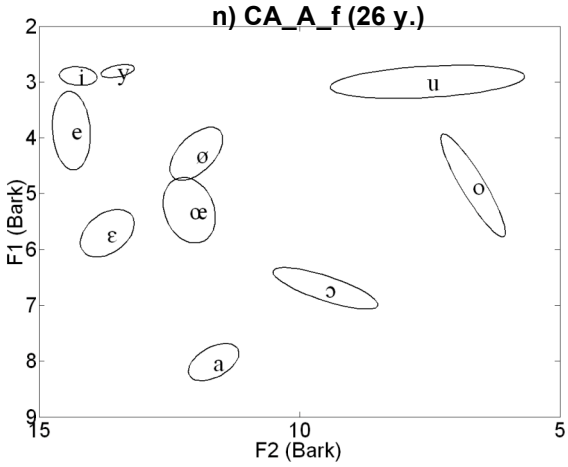

i) CA_8_f (8 y. 4 mo.)
F1 (Bark)
F2 (Bark)
i
y
u
e
ø
ɛ
œ
o
ɔ
a


j) CA_8_f (9 y. 1 mo.)
F1 (Bark)
F2 (Bark)
i
y
u
e
ø
ɛ
œ
o
ɔ
a


k) CA_A_f (22 y.)
F1 (Bark)
F2 (Bark)
i
e
y
ø
u
o
ɛ
œ
ɔ
a


l) CA_A_f (22 y.)
F1 (Bark)
F2 (Bark)
i
y
u
e
ø
o
ɛ
œ
ɔ
a


m) CA_A_m (23 y.)
F1 (Bark)
F2 (Bark)
i
y
u
e
ø
o
ɛ
œ
ɔ
a


n) CA_A_f (26 y.)
F1 (Bark)
F2 (Bark)
i
y
u
e
ø
o
œ
ɛ
ɔ
a

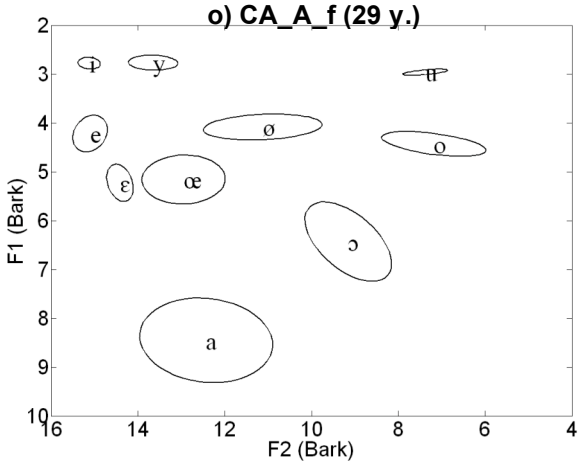

o) CA_A_f (29 y.)
F1 (Bark)
F2 (Bark)
ɨ
y
ʉ
e
ø
o
ɛ
œ
ɔ
a

FIGURE 4:

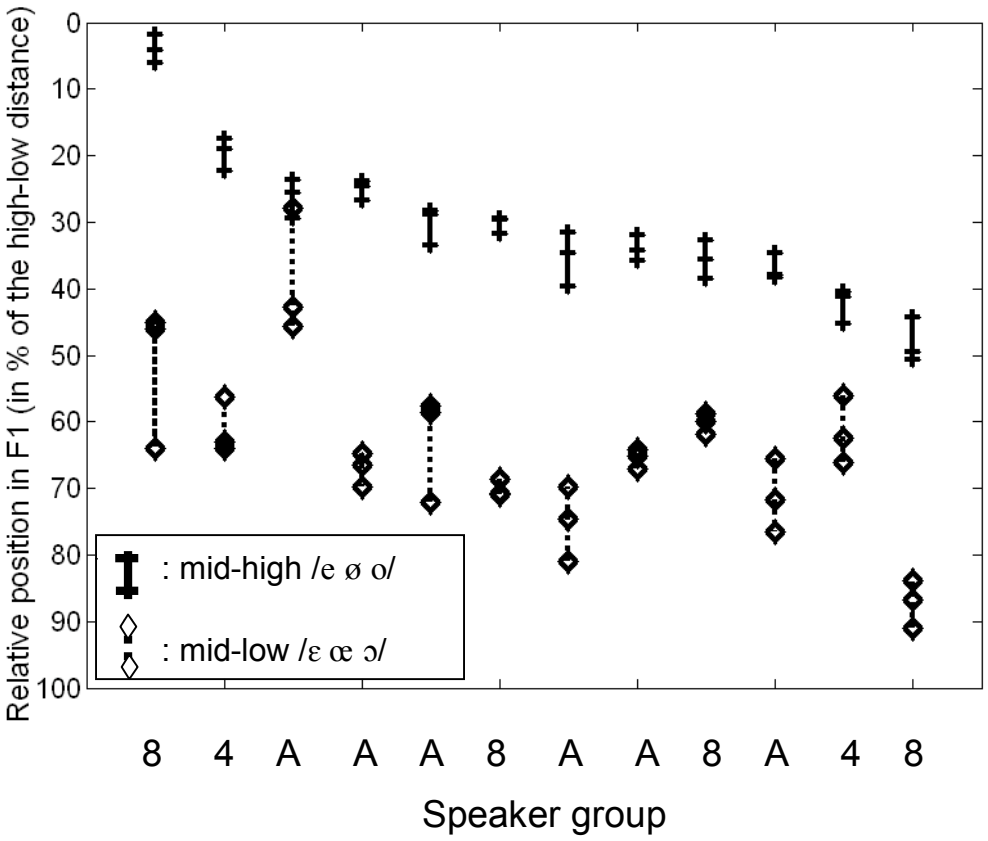

Relative position in F1 (in % of the high-low distance)
0
10
20
30
40
50
60
70
80
90
100
: mid-high /e ø o/
: mid-low /ɛ œ ɔ/
8 4 A A A 8 A A 8 A 4 8
Speaker group

FIGURE 5:

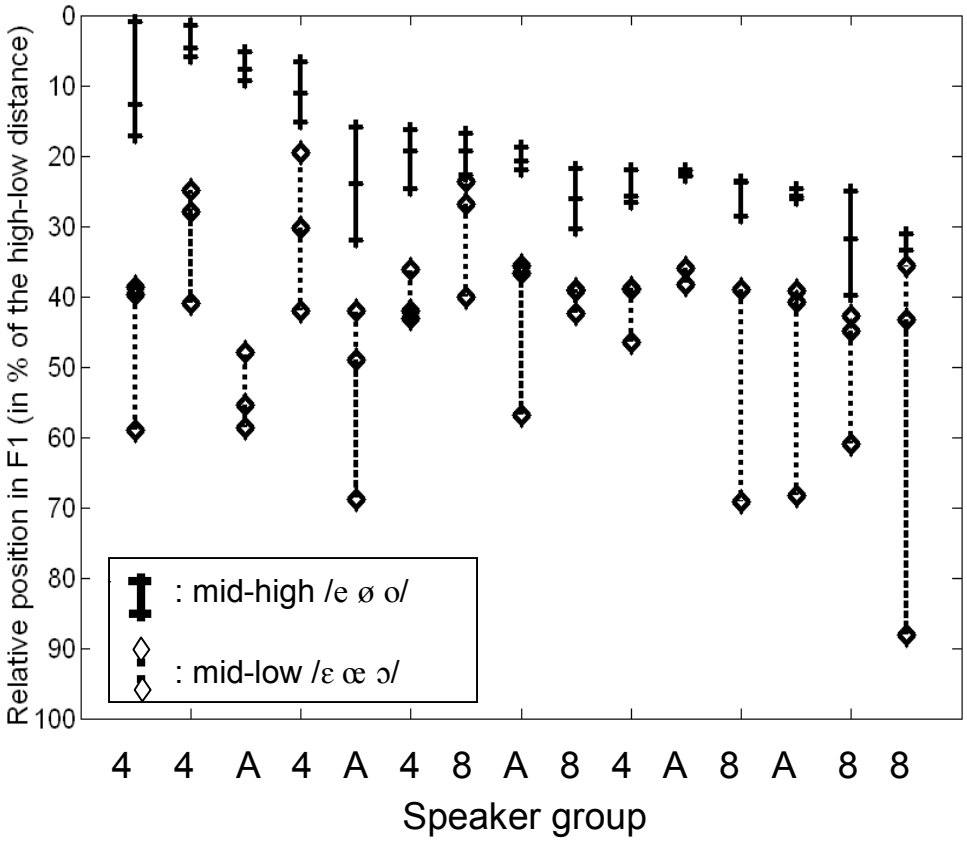

Relative position in F1 (in % of the high-low distance)
0
10
20
30
40
50
60
70
80
90
100
: mid-high /e ø o/
: mid-low /ɛ œ ɔ/
4 4 A 4 A 4 8 A 8 4 A 8 A 8 8
Speaker group

FIGURE 6:

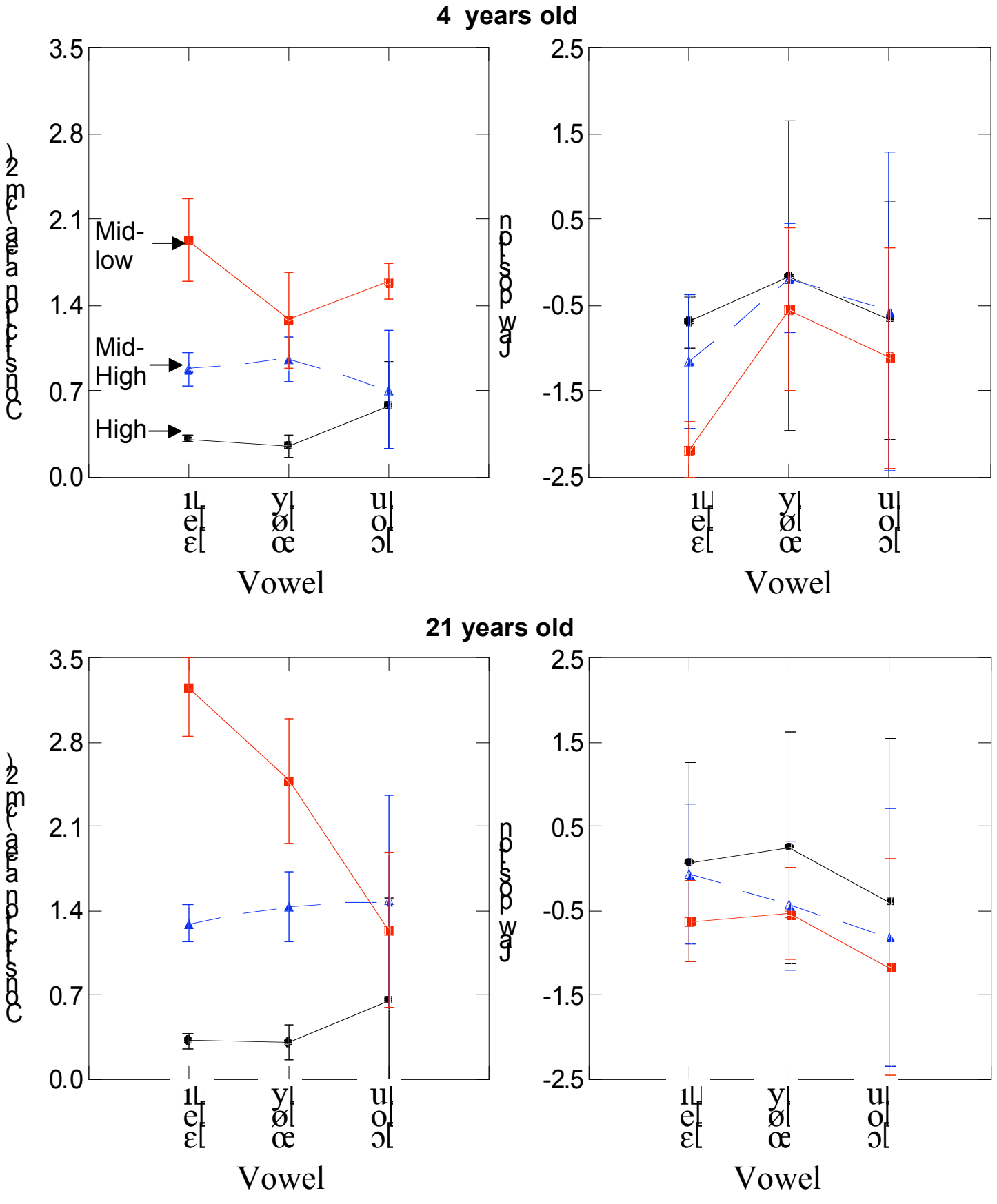
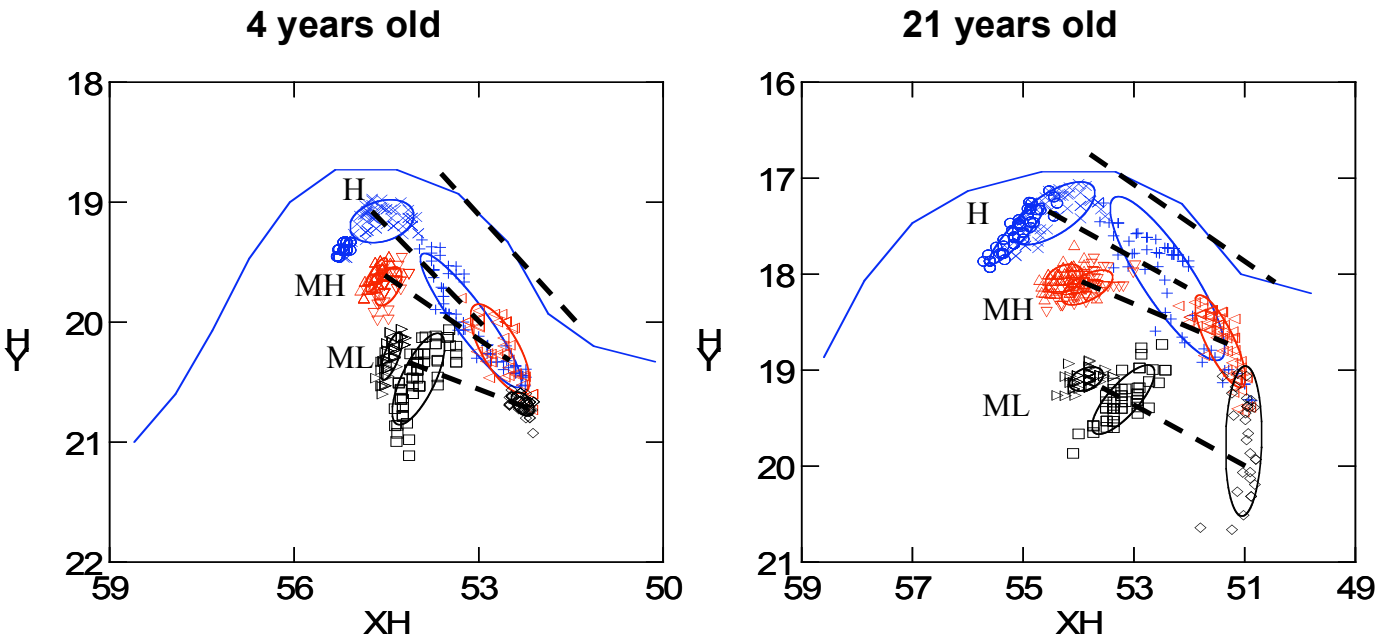

4 years old
21 years old
Mid-low
Mid-High
High
3.5
2.8
2.1
1.4
0.7
0.0
2.5
1.5
0.5
-0.5
-1.5
-2.5
ɪ ɛ e
y ø œ
u o ɔ
Vowel

Figure 7:


4 years old
21 years old
H
MH
ML
XH
18
19
20
21
22
16
17
59
56
53
50
57
55
51
49