Acoustic aspects of vowel harmony in French

2 authors:

Noël Nguyen
Aix-Marseille Université
106 PUBLICATIONS **1,016** CITATIONS

SEE PROFILE

Zsuzsanna Fagyal
University of Illinois, Urbana-Champaign
52 PUBLICATIONS **460** CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Project COBRA (Conversational Brains) View project

Project Rhythm in working-class Parisian French View project

Acoustic aspects of vowel harmony in French

Noël Nguyen (1) and Zsuzsanna Fagyal (2)

(1) Laboratoire Parole et Langage, Aix-Marseille Université & CNRS
29 avenue Robert Schuman, 13621 Aix en Provence, France

(2) Department of French, University of Illinois at Urbana-Champaign
2090 FLB, 707 S. Mathews Avenue, Urbana, IL 61801, USA

running title: Acoustic aspects of vowel harmony in French

Abstract

This paper examines acoustic aspects of vowel harmony (VH), understood as regressive vowel-to-vowel assimilation, in two regional varieties of French in six speakers' productions of 107 disyllabic word pairs. In each word pair, the word-initial vowel (V1) was phonemically either /e/ or /o/, and the word-final stressed vowel (V2) alternated between /e-ɛ/, /ø-œ/, /o-ɔ/ or /i-a/. Results are consistent with the idea that VH in French entails variations in tongue height along with related displacements of the tongue position along the front-back axis. These effects were independent of both the number of morphemes and lexical frequency. They were more systematic in Northern than in Southern French speakers' speech. Linear mixed-effects models strongly suggest that VH is a gradient effect of the trigger on the harmonizing vowel. Results lend support to usage-based phonological approaches regarding gradient phonetic differences as part of the gestural scores that make up the lexicon and that can be variably grammaticalized in different varieties of the language.

1 Introduction

1.1 From coarticulation to vowel harmony

It has been suggested that vowel harmony (henceforth, VH), a morpho-phonological process in many languages requiring all vowels in a word to share or to "harmonize in" one or several features, could be understood as the "grammaticalized" end product of "an earlier phonetic process involving vowel-to-vowel assimilation" (Ohala, 1994). According to this hypothesis, VH arose gradually through sound change due to a dissociation parsing error: "the perturbation on a vowel caused by an adjacent vowel" (i.e. vowel-to-vowel assimilation triggered by a vowel) became so strong as to cause listeners to parse the harmonizing vowel as independent of the triggering vowel. Once dissociated from its trigger, the quality of the harmonizing vowel could vary independently from its initial conditioning environment, and thus be assigned new functions in the grammar, i.e. develop into a separate morpheme.

While several studies showed that vowel-to-vowel assimilation processes are indeed common to many languages (Öhman, 1966), while it is even proposed that the direction of assimilatory processes would show a universal anticipatory rather than perseverative bias (Hyman, 2002), the possible historical link between VH and vowel-to-vowel assimilation, a fascinating hypothesis, remains little understood.

In this article, we examine whether vowel-to-vowel assimilation exists, and if so, whether it can be manifest to varying degrees in different varieties of French. Based on the acoustic examination of 107 disyllabic word pairs inserted in carrier sentences in two regional varieties of French, we investigate whether vowel-to-vowel assimilation—traditionally referred to as VH—is more or less systematic in the Northern rather than the Southern regional dialect of French spoken in France. We hypothesize that spoken varieties of the same language could, indeed, be aligned on a continuum stretching from inconsistent and marginal to systematic and near-phonemic alternations in vowel height depending on the quality of an adjacent stressed vowel. We expect that variations in the "strength" or "consistency" of VH-like assimilatory phenomena in French depend to a large

3

extent on the speaker and the type of the vowel (e.g. back vowels reacting to height harmony differently from front vowels), as well as on the speaker's local dialect.

We suppose, following Ohala (1994), that more or less strongly variable acoustic realizations of VH similar to patterns of variation in our data could lead to variable parsing of the phonetic input by listeners, which could ultimately result in gradual dissociation and phonological reanalysis of the quality of the harmonizing vowel. We are thus not only concerned with "extracting generalizations about patterns of VH and explaining them by reference to universal phonetic factors" (Ohala, 1994), but also with studying both internal morpho-lexical and external speaker-specific factors, such as morpheme count, lexical frequency, and local dialect, which could play a role in triggering or reinforcing changes in listeners' variable parsing of the phonetic input.

1.2   Vowel harmony in French

Compared with other vowels, mid vowels in French are known to show large variations in quality. This variability has been ascribed to several factors. One of these factors is called the *loi de position* ("law of position", or closed/open syllable adjustment), which refers to the fact that mid vowels tend to be lower in closed syllables and higher in open syllables, e.g. /pø/ *peu* "little/few" vs /pœʁ/ *peur* "fear". Other sources of variation in the realization of mid vowels include the identity of the post-vocalic consonant, the morphophonological structure of the word, and vowel harmony. VH in French is a type of regressive assimilation which causes the vowel of a stressed syllable to affect the pronunciation of a mid vowel in a preceding open syllable. In the simplest terms, it entails that the vowel preceding the stressed vowel tends to be mid-high when the stressed vowel is mid-high or high (e.g. *aimer* [eme] "to love"), and mid-low when the stressed vowel is mid-low or low (*aimable* [ɛmabl] "kind", see Walker, 2001). Assimilation is only partial, shows considerable variability, and seems principally to affect vowel height. Variations along the front-back dimension may also occur insofar as they are correlated with vowel height (Grammont, 1939; thus, one may assume that the harmonizing vowel is slightly retracted in *aimable* relative to *aimer*).

Previous studies identified multiple sources of possible phonetic and external variation in real-

4

izations of VH. The following brief review of the literature will highlight most of these factors, bearing in mind that what is known of VH in French prior to our study has been predominantly based on self-reports or impressionistic analyses of words or word pairs in citation forms. In these contexts, the final primary stressed vowel (henceforth V2) of the bisyllabic word is considered the trigger and the preceding word-initial vowel (henceforth V1) the harmonizing unstressed vowel[1]. VH in French is often said to be restricted to contexts in which V1 is a mid vowel. According to Casagrande (1984), for instance, VH only affects mid vowels: therefore the mid vowel /e/ in the word *bêtise* cannot be "turned into a high vowel" (p. 90), i.e. raised to the extent of sounding more like an /i/[2]. For most authors, VH is restricted to the front unrounded pair /e-ɛ/ (Dell, 1973; Fouché, 1959; Malmberg, 1969), although Grammont (1939) extends it to /œ/ (see also Casagrande, 1984), and Tranel (1987) to both /œ/ and /ɔ/ when these are followed by their mid-high counterparts /ø/ and /o/.

In many studies, VH is considered to be solely a raising phenomenon such that mid-low vowels are raised when followed by a high, or mid-high vowel. Thus, as Casagrande (1984) notes, "the tendency is always toward closure; that is, no vowels harmonize by causing a wider opening of the mouth" (see also Fouché, 1959; Grammont, 1939; Malmberg, 1969; Tranel, 1987). The set of possible triggering vowels is often further confined to front vowels, and more specifically to /i, y, e/ (Casagrande, 1984; Fouché, 1959; Grammont, 1939; Malmberg, 1974). A noticeable exception is Dell's (1973) HARM rule according to which /e/ can be realized as [ɛ] before a low vowel (see also Walker, 2001).

The role of spelling in realizations of VH has also received relatively little emphasis. With the exception of Fouché (1959) who was the first to invoke the possible effect of spelling, Tranel (1987) is almost alone suggesting that some orthographic representations might be directly tied to phonological representations: for instance, in the case of the letter *é* that could trigger a more closed realization of the front-mid vowel, whereas the letters *è*, *ê*, *ai*, and *e* before two written consonants may favor a more open pronunciation of the mid vowel. (This aspect is not examined in this paper.)

5

The possible influence of syllable structure, namely the presence and type of a coda consonant following V1, on the strength or presence of VH, has also been little studied. Although this otherwise important aspect of VH is not investigated here, it is noteworthy that some, e.g. Fouché (1959), rejected the possibility that front mid vowels in closed syllables before /ʁ/, e.g. in *percer* "to drill" vs. *(il) perce* "(he) drills", could undergo VH. The number of segments between the target (V1) and triggering (V2) vowels as a possible factor in VH surfaces in Dell's (1973) discussion of the case of words in future tense, e.g. *(vous) aiderez* "(you) will help" which contain an underlying schwa that is supposed to act as a barrier to VH.

As for the morpho-phonological make-up of the word, VH is usually characterized by comparing the base form of a word with a form containing a derivational or inflectional suffix. In most examples quoted in previous studies, the base form (which Grammont, 1914, referred to as "le mot plus simple et plus court", the simpler and shorter word) is a monosyllable with a mid-low or low vowel, e.g. *aigre* /ɛgʁ/ "bitter" compared with *aigri* /egʁi/ "embittered" (Fouché, 1959; Walker, 2001), *prêt* /pʁɛ/ "a loan" compared with *prêter* /pʁete/ "to lend" (Grammont, 1939).

It is a widely shared assumption that VH operates between the vowels in the suffix and root, in the derived or inflected form. Dell (1973) considers that the presence of a morpheme boundary between the two vowels is in fact necessary for VH to occur, i.e. that VH does not happen within morphemes (hence, there would be only one possible phonetic realization for *récolte*, one morpheme, "harvest", namely [ʁekɔlt], as opposed to *[ʁɛkɔlt], in spite of the second vowel being mid-low). Tranel (1987), conversely, suggests that VH also applies to monomorphemic words (e.g. *auto*, one morpheme, "car", realized as [oto]).

Despite the fact that the majority of phonetic and phonological treatises point towards the existence of VH as an important characteristic of French, almost all of them also consider it optional. They point out that it is very likely speaker-dependent and more easily perceptible in conversation than in formal speech contexts (Coveney, 2001; Dell, 1973; Fouché, 1959; Walker, 2001).

Although Grammont's (1939) study, the first study on VH in French, explicitly attributes VH to "Parisian French", there is virtually no information available on the potential effect of regional

6

accent on patterns of VH. Most authors confine the scope of their study to so-called Standard or "General" French[3].

Thus with a few modifications, the majority of studies and treatises that recognize the existence of VH-type phenomena in French consider it an anticipatory process affecting primarily front-mid vowels. But among the strong advocates of VH, there are also those who are sceptical of its importance. Functional phonological approaches seem to belong to the latter. Martinet (1945) affirmed having found no evidence that the first vowel in a lexical word would be affected by the quality of the following vowel. Based on questionnaires circulated in a POW camp for French officers, Martinet declared that in none of the regional pronunciations of French does the assimilation process seem to be typical of "the majority of the subjects" (p. 142). Therefore, he argued that such a phenomenon "does not seem to present the universal character many phoneticians, including Mr. Grammont, wish to attribute to it" (idem). It is doubtful, however, whether Martinet's definition of *dilation*, i.e. the term coined by Grammont (1939) for what is known and referred to as VH today, corresponds to Grammont's. Upon closer inspection it turns out that the test cases Martinet presented to his subjects are not discussed in, and sometimes not even related to, cases of vocalic dilation as defined by Grammont. VH transforming the high front mid vowel /e/ into its open counterpart /ɛ/, for instance in the *été* /ete/ "summer" vs. *était* /ɛtɛ/ "(s/he) was" pair, presented as part of Martinet's word list, is not discussed by Grammont, and the presumed alternation of the rounded mid vowels /œ/ and /ø/, exemplified by Martinet's *déjeuner* /deʒøne/ vs. *déjeunons* /deʒœnõ/ example, is not a type of alternation taken into account by Grammont's dilation rule.

Landick (1995) took an essentially functional phonological approach to the question of VH when she perceptually analyzed the production of numerous target words in sentences and citation forms uttered by native speakers of Parisian French of different social categories. Her conclusions pointed towards a marginal role of VH in contemporary French. In a more recent book-long treatise, however, Landick (2004) revisits her conclusions. Besides additional details about the impressionistic analyses she conducted, she adds as a novelty some preliminary acoustic analyses of

7

citation forms of word pairs (*bégaie-bégaiement*) and cohorts (*cloche, clocher, clochette, clochard*) and concludes that there is acoustic evidence of VH in her production data, even though the pattern is highly variable and speaker-dependent.

Although VH is mentioned in most textbooks on the phonetics and phonology of French, apart from suggestive illustrations (Boula de Mareüil and Fagyal, 2000; Landick, 2004) and a tentative first look at acoustic correlates (Fagyal et al., 2003), its articulatory and acoustic aspects have never been systematically explored. The goal of the present paper is to offer the first such study. Its aim is to identify the acoustic correlates of VH by characterizing its effect on the spectral shape of non word-final front and back mid vowels elicited in carrier sentences from several native speakers from two different dialect areas in France. As well as the influence of the dialect, the effects of morphological complexity and lexical frequency will be assessed.

2   Method

2.1   Material

The corpus consisted of 107 pairs of disyllabic words, drawn from three word classes: nouns (common and proper names), adjectives, and infinitives. The first syllable of each word of each pair contained a mid vowel (henceforth, V1) which was phonemically identical in both words of the pair. This vowel was either front and unrounded or back and rounded, depending on the word pair. The second syllable contained a vowel (henceforth, V2) which was, in most pairs, non-low in one word and low in the other word of the pair. For part of the word pairs, however, there were differences in the distribution of non-low and low vowels in Syllable 2 depending on the speaker's dialect, which are discussed below. The onset of the first syllable, if filled, consisted of a single consonant, or an obstruent+liquid cluster. The onset of the second syllable was either a single consonant or an obstruent+liquid cluster. The first syllable was open and the intervocalic consonant or consonant sequence affiliated to the second syllable in all of the pairs.

Table I shows the sets of word pairs with one example for each set, the frontness/backness of V1, the expected difference in the phonetic form of V2 between the two words, and the total number

of word pairs in each set[4].

V1s are systematically transcribed in Table I as high-mid vowels. This is consistent with the *loi de position* (see Section 1.2), which we adopted as a rule of transcription. The actual quality of V1 was of course an empirical issue that this acoustic study was precisely intended to address. When V2 was a mid vowel, as was the case in the first six word-pair sets contained in Table I, its expected phonetic form was determined as follows: 1) In the *été-éther* and *potée-poterne* sets, the second syllable was always open in the first word and closed in the second word, and the realization of V2 in those words generally conformed to the *loi de position*; 2) in the *prêteuse-prêteur* and *poseuse-poseur* sets, V2 was always followed by /z/ in the first word and by /ʁ/ in the second word, and its phonetic realization was expected to reflect the complementary distribution of [ø] and [œ] before /z/ and /ʁ/ in word-final syllables in Standard French; 3) in the *dévot-dévote* and *auto-automne* sets, the second syllable was open in the first word and closed in the second word in all but one pair (*épaule-époque*), and the realization of V2 was expected to be governed by the *loi de position* again. In *épaule-époque*, V2 is also expected to be high-mid in the first word and low-mid in the second word in Standard or General French (see Note 3). In the last two word-pair sets in Table I, we expected V2 to be systematically pronounced as a high vowel in the first word and as a low vowel in the second word of each pair. The speakers' actual realizations, as we will see however, did not necessarily conform to these expectations.

2.2   Speakers and recordings

Six speakers were recorded, three speakers each from two main dialect areas in France (henceforth, North and South)[5]. Biographic information is provided in Table II. The speakers' age and educational levels were comparable: all of them were in their twenties or thirties at the time of the recording and were college educated. All of them volunteered to participate in the experiment. Speakers S1-S3 were born and/or grew up in or around Paris, where they lived until at least the end of their undergraduate education. Their colloquial and formal reading styles represented the

Northern French variety. Speakers S4-S6 were born in the area of Aix en Provence. They lived at their place of birth until at least the end of their secondary education and are currently living in Aix en Provence. Their colloquial French represents the Southern French variety. During the recordings, however, we could perceptually identify S4 and S5's tendency to style-shift, i.e. adopt a formal reading style that shared several features with General French, in that it tended to have mid-high vowels as opposed to mid-low vowels before a tautosyllabic /z/, and did not have post-consonantal word-final schwa, e.g. *prêteuse* [pʀ̥et̪øz]. In contrast, S6 seemed to have both mid-low vowels prior to a tautosyllabic /z/ and word-final schwa, generally realized as a front rounded vowel, e.g. [pʀ̥et̪œzø]. Thus, while V2 gave the auditory impression of a systematic contrast in height within each word pair for the other speakers, consistent with the transcriptions in Table I, this contrast sounded neutralized for S6 in the *prêteuse-prêteur* and *poseuse-poseur* sets (27 pairs).

— Insert Table II about here —

Speakers S1 and S3, enrolled as graduate and postgraduate students at UIUC, were recorded in the UIUC Phonetics Laboratory. Speaker S2's recording took place at the Phonetics Institute of Paris. Speakers S4, S5 and S6 were recorded in the Laboratory for Speech and Language in Aix en Provence. At all three locations, the recordings were carried out in an anechoic chamber using a high-quality microphone and a DAT recorder. The material was presented to the speaker on slides shown on a computer screen. Each word appeared twice on each slide, in a carrier sentence as in the following example: "Il retape *été* parfois; *été*" /il ʀətap ete paʀfwa ‖ ete/ "he retypes *summer* sometimes; *summer*".

The sentence referred to a man typing words on a keyboard. The first instance of the target word appeared between two symmetrical phonetic sequences (/ap/ and /pa/). The bilabial stops to the left and right of the utterance-medial target word were expected to have a minimal influence on the shape of the tongue in the production of that word. The utterance-final second renditions of the target words were intended both as stimuli in a companion perceptual experiment (Nguyen, Fagyal & Cole, 2004) and as additional material for this study. The speakers were asked to read each sentence at a normal rate, with no special emphasis on any word, and no pause between

10

words in the first intonational phrase (up to *parfois* "sometimes"). The utterance-final rendition of the target word had to be as phonetically similar to the first one as possible. Speakers were also instructed to maintain the same speech rate, rhythm, and intonation contour throughout their reading of the corpus. Short pauses were inserted every thirty sentences.

The list of sentences was recorded by each speaker twice, in two separate sessions, with at least a week interval between the sessions. Sentences were presented in a different random order in each session. The data therefore included four repetitions of each word for each speaker. They were transferred onto the hard disk of a computer (sampling rate: 22050 Hz, 16-bit integers) for future processing.

2.3 Acoustic measures

The vowels in each target word were manually segmented from the waveform and the corresponding wide-band spectrogram. Markers were placed at the acoustic onset and offset of the first and second vowels. The frequencies of f_0, F_1 and F_2 were automatically measured every 5 ms over the entire utterance using the Entropic ESPS library. The settings of the formant tracker were the following: low-pass filtering at 10 kHz; preemphasis constant: 0.94; 49-ms cos^4 window; LPC order: 12. f_0, F_1 and F_2 frequencies at the acoustic midpoint of the first vowel and that of the second vowel were automatically extracted. Formant frequency values were then verified manually. For each speaker and each vowel in each V1-V2 combination, extreme frequency values were checked using both an FFT spectrum and an LPC spectrum computed over a 49-ms window centered at the midpoint of the vowel, together with a wide-band spectrogram, and they were corrected when it was found that the formant had not been correctly located by the automatic formant tracker. Overall, 838 (3.7 %) out of 22848 formant frequency values were remeasured by hand in this way, and 653 (2.8 %) of them were corrected.

There were, as expected, differences in the number of detected automatic measurement errors depending on the phonetic identity and position of the vowel in the word. Errors were overall more numerous for vowels in the first than in the second syllable, all vowels and speakers combined ($N =$

11

413 (3.6 %) in V1 and $N = 240$ (2.1 %) in V2). Errors were found to be more numerous for /o/ in the first syllable (V1) than for the other vowels both in the first and second syllables (positions V1 and V2)[6]. The number of detected errors differed depending on the speaker. Overall, automatic measurements contained more detectable errors for female speakers than for male speakers for all vowels in the second syllable ($N = 212$ (3.7 %) vs. $N = 28$ (0.5 %)), while in all vowels combined in the first syllable female and male voices roughly showed the same error rate ($N = 196$ (3.4 %) vs. $N = 217$ (3.8 %)). The error rate was higher for the female speaker S4 than for the other speakers. One potential origin for these errors was S4's breathy voice, which resulted in a high-amplitude fundamental component in the spectrum, and which may further be associated with additional spectral peaks caused by tracheal coupling (Klatt and Klatt, 1990).

In the next step, the linear formant frequency scale was converted to a Bark scale using the following formula (Traunmüller, 1990):

$$z(F_j) = 26.81 \ / \ (1 + (1960 \ / \ F_j)) - 0.53, \qquad j \in \{1, 2\}$$

where F_j represents the frequency in Hz of the j-th formant and $z(F_j)$ the corresponding value in Bark[7].

$z(F_1)$ and $z(F_2)$ were taken as acoustic/auditory correlates of the vowel-height and front-back dimensions, respectively, as these dimensions are traditionally defined[8]. Statistical analyses were performed on the Bark-transformed formant frequencies. For the sake of simplicity and familiarity for the reader, however, frequency values presented in the remainder of this paper will be given in Hertz. These values were derived from Bark values using the inverse transform $f = 1960/[(26.81/(z + 0.53)) - 1]$. In addition, we will refer to $z(F_1)$ as F_1 and to $z(F_2)$ as F_2.

3   Results

3.1   Acoustic properties of the second vowel

Our analyses first focused on the acoustic properties of V2, the stressed vowel and expected trigger of VH. We sought to characterize variations in the formant pattern of V2 with a view to using them as a reference frame with respect to which V2-dependent differences in V1 would be interpreted. We also aimed to identify potential differences in the phonetic realization of V2 between speakers, and to predict the impact that these differences may have on V2-to-V1 regressive harmony patterns. The distribution of V2 in the F_1 vs F_2 plane is shown in Figure 1 for each speaker separately.

— Insert Figure 1 about here —

Figure 1 indicates that, unsurprisingly, front vowels are situated along a slanted straight line inclined to the left, reflecting the fact that the mid-low vowel /ɛ/ is more retracted than its mid-high counterpart and, of course, than the vowel /i/. The relative distribution of these vowels with respect to each other, however, shows important speaker-dependent variations. Contrary to expectations (see Mettas, 1979; Coveney, 2001), the open vowel /a/ did not register as a front vowel within the individual vowel space of Northern speakers S1, S2 and S3, with the noticeable exception of Speaker S2, whose realizations suggest lesser height (F_1 dimension) and more fronting (F_2 dimension) of /a/ than it is the case for the other speakers. (This male speaker was born and has lived all his life in Paris.)

The relative location of back vowels displays the greatest amount of variation between speakers. As expected, /o/ is more retracted than /ɔ/ in each case, but the degree of backness of both vowels varies greatly relative to /a/ from speaker to speaker. Two female speakers, one from the North (S1) and one from the South (S6) exhibit the most extreme differences, while the overall vowel patterns for the other speakers can be regarded as taking place in-between these two endpoints. In S1, compared with all other speakers, /ɔ/ and to a lesser degree /o/ appear to be considerably fronted, with F_2 systematically greater than 1000 Hz. The vowel space of S6, on the other hand,

13

is organized very differently, and reflects a more canonical distribution of back vowels. In her case, the two back vowels and /a/ can be viewed as located on a slanted straight line, indicating that /o/ and /ɔ/ are both more retracted than /a/, given as a central vowel in most dialects of French. The two other speakers whose back vowel patterns are similar to that for S6 are also Southern speakers. The back vowel space of S5 is probably closer to the pattern shown by S6 than S4, a female speaker, representing an intermediate pattern between those of S1 and S6. The two other Northern speakers (S2 and S3) also display a clear tendency for fronting: S3 slightly more than S2.

Although this study had not set out to explore in-depth, dialect-dependent differences in the articulation of V2 (and the preceding visual inspection of the speakers' vowel spaces is no substitute for such study), such dialectal differences clearly seem to emerge from our data. The fronting of /o/ and /ɔ/ is a well-established trend in Northern French (see Malderez, 1995; Coveney, 2001), of which Speaker 1 was impressionistically judged as a good representative. Such fronting has not been reported in Southern varieties of French, of which S6's speech forms the most typical example in this study.

In Southern French varieties, there is no contrastive distinction between /e/-/ɛ/, /ø/-/œ/ and /o/-/ɔ/, and the distribution of the mid-high and mid-low variants is said to be entirely governed by a variant of the *loi de position*: the mid-high allophone occurs in open syllables and the mid-low allophone in closed syllables and whenever the next syllable contains a schwa (Durand, 1990). On this account, V2 is therefore assumed to be mid-low in both words of pairs such as *prêteuse-prêteur* and *poseuse-poseur*. Figure 1 shows that Speaker S6 tended to conform to this pattern as the vowel transcribed /ø/ in standard French is much closer to /œ/ for her than for the other speakers. Accordingly, we expected V2-dependent variations in the acoustic shape of V1 to be minimal in the *prêteuse - prêteur* and *poseuse - poseur* word sets for Speaker S6.

14

3.2 Vowel harmony patterns

Variations in the formant pattern of V1 depending on the second vowel are schematized in Figures 2-7 for the different speakers. Each figure contains four panels corresponding to the four V2 pairs. Each of these panels illustrates the variations shown by the vowel /e/ in V1 (henceforth, front V1), to the left, and by the vowel /o/ in V1 (henceforth, back V1), to the right, as a function of V2. Open and filled circles represent average formant frequency values for V1 before low and non-low vowels, respectively. The formant frequency values for /i, e, ø, ɛ, œ, a, ɔ, o/ in V2 position, averaged across all the words in which these vowels appeared in our material, are also shown as a reference grid in grey[9].

— Insert Figures 2-7 about here —

We begin with a few observations on the overall characteristics of V1, regardless of the V2-related variations the vowel may display. V1 generally appears to be closer to a mid-high than to a mid-low vowel. This is consistent with the *loi de position* that predicts high-mid vowels in open syllables, as well as with our transcriptions in Table I, even though this tends to be true for the front V1 /e/ to a greater extent than for the back V1 /o/. A noticeable exception is Speaker S2, whose back V1 shows the opposite trend and is overall closer to mid-low than to mid-high. Another general characteristic is that the back V1 is relatively fronted for this speaker. Such a tendency mirrors the fronting of back vowels in V2 position as described in the preceding section. As for V2, it is contingent upon the speaker's dialect: it applies to the Northern speakers (S1-S3) and to one Southern speaker (S5), whereas the two other Southern speakers, and more characteristically S6, do not exhibit this tendency.

As indicated in the introduction, vowel harmony in French is traditionally regarded as a phenomenon of regressive assimilation in height. Figures 2-7 suggest that V1 tends to show variations in the expected direction. In general, it appears to be higher in the context of a mid-high or high as opposed to mid-low or low following vowel. In addition to height distinctions, Figures 2-7 provide indications that variations along the front-back dimension tend to occur for V1 depending on V2.

15

These indications are consistent with front V1 being fronter when the following vowel is mid-high or high than when it is mid-low or low. Back V1 is characterized by two distinct patterns of variation. On the one hand, Figures 2-7 suggest that it is more retracted before a mid-high vowel (more specifically, /e/ or /o/) than before a mid-low vowel (/ɛ/ or /ɔ/). On the other hand, back V1 appears to be fronter before the high vowel /i/ than before the low vowel /a/. The behavior of back V1 prior to /ø/-/œ/ was less systematic. Overall, the observed vowel-harmony effects are both relatively subtle and more consistent for Northern than for Southern speakers.

A series of repeated-measure ANOVAs was performed to assess the statistical significance of the above tendencies. An ANOVA was conducted for each V1 × V2-pair combination separately for each speaker. In each analysis, the word pairs were used as a random variable and V2's height as the independent variable. No correction procedure was applied across the ANOVAs.

Mean differences in F_1 and in F_2 for V1 depending on V2 for each speaker are shown in Table III. A positive difference indicates that the value of the parameter of interest is higher when V2 is mid-high or high compared with mid-low or low. By way of example, F_2 was found to be higher by 73 Hz on average for the front V1 before /e/ compared with /ɛ/ for Speaker S1 (second row from top, leftmost column). Statistically significant differences ($p < 0.05$) are in bold. The associated F and p values are given in Table IV.

— Insert Table III about here —

— Insert Table IV about here —

Table III shows that F_1 is generally lower in frequency before a mid-high or high vowel than before a mid-low or low vowel. This is systematic prior to /i/-/a/, especially for the front V1, in which case the difference is statistically significant for all speakers. This is also consistently observed before /ø/-/œ/ in particular for the back V1 (*poseuse-poseur* word set) in Northern speakers, and before /e/-/ɛ/ for Speakers S2, S3 and S5. The lowering of F_1 before mid-high or high vs. mid-low or low vowels applies to all Northern speakers, and to a lesser extent to Southern speakers, especially Speaker S6, the most prototypical "Southern" speaker in the corpus, who tends to

16

deviate from this trend. Note, however, that in none of the cases is the F_1 difference significantly positive.

Table III also indicates that for the front V1, F_2 is in general higher in frequency when V2 is mid-high or high compared with mid-low or low, with a few exceptions (non-significant negative differences before /ø/-/œ/ for S4 and S5 and before /o/-/ɔ/ for S4). This tendency is particularly systematic for Northern Speakers. For the back V1, F_2 frequency is consistently lower before a mid-high than a mid-low vowel (with the exception of one non-significant positive difference for S6 before /ø/-/œ/). Conversely, it is larger before /i/ than before /a/ (always significantly so except for S6).

Common characteristics between front and back V1s emerge when these vowels are examined separately before mid and non-mid vowels (/i/ or /a/). First, V1, whether front or back, is higher when followed by a high-mid compared with a low-mid vowel. For front V1s, an increase in height is associated with a lower F_1 frequency and a higher F_2 frequency. For back V1s, it is associated with both a lower F_1 and F_2 frequency. (The covariation between F_1 and F_2 reflects the fact that, as already pointed out in Section 3.1, both front and back vowels are located along slanted lines in the F_1-F_2 plane, with, from bottom to top, a leftward slope for the front vowels, and a rightward slope for the back vowels.) Second, our data suggest that V1 is consistently higher and more fronted before /i/ than before /a/. In this latter case, the pattern of variation of V1 may clearly be ascribed to a mechanism of articulatory accommodation to the upcoming vowel. However, the mechanism responsible for the variations of V1 prior to a mid vowel may be of a different nature. We will return to this in the general discussion.

3.3 Potential influence of morphological structure on vowel harmony

As indicated in the introduction, it has been assumed that vowel harmony in French "respects the morphemic cut of words" (Dell, 1973) in the sense that VH would be appropriate in bi-morphemic contexts such as /sed+e/ (*céder* "to yield"), but would not operate in mono-morphemic words such as *récolte* "harvest" and *prétend* "pretends". In this section, we examine whether VH varies

17

in magnitude depending on the presence or absence of a morpheme boundary between the two vowels.

Although the presence and number of morpheme boundaries can be straightforwardly determined in nouns and adjectives whose forms alternate in gender (e.g. *prêteuse-prêteur*), the morpheme count varies depending on the morphosyntactic category and the etymological transparency of words. The word *défaite* "defeated" (adj. fem.) is considered tri-morphemic if it is an adjective, since the pronounced final consonant /t/ in the spoken form (the feminine *e* in the written form) is considered an alternating morpheme that encodes gender distinctions: *dé-* (prefix), *-fait-* (base), and /t/ or *-e* (suffix). However, when *défaite* "defeat" is a noun, the word is considered bimorphemic: *dé-* (prefix), *-faite* (base). While the derivational suffix *-é* in noun-based adjectives such as *fléché* is undisputedly an independent morpheme, what is perceived today by many as the suffix *-aire* added to the verbal base *noter* in the word *notaire* "notary" is not a suffix in light of the originally borrowed, but now monomorphemic Latin form *notare*. Opinions about the status of a morpheme are also split when the word is a recent borrowing. The word *motel*, initially borrowed from the compound "motor hotel" in English, has lost or never retained its original bi-morphemic status, and therefore it is now considered to be mono-morphemic in French.

These uncertainties about the exact morphemic cut of target words in our corpus led us to ask three linguists, experienced in phonological and lexicographic analyses of French, to determine the number of morphemes in the spoken forms of the target words used in this study. These three coders were given the task of deciding whether a word was monomorphemic or bimorphemic. They could also decide that the word contained more than two morphemes in which case they had to justify their choice with comments. In case of uncertainties about the meaning of a word, the coders were invited to consult dictionaries.

The majority of the words in the corpus were coded as bimorphemic by at least two of the three coders: 60% ($N = 82$) of the first words of each pair (henceforth, W1s) and 61% ($N = 83$) of the second words of each pair (henceforth, W2s) were considered bimorphemic, while the rest of the words were coded as monomorphemic. Two coders in five separate occasions suggested

a trimorphemic analysis, for *défaite*, *méfaire*, *fléchage*, *réglage*, and *collage*. Thus the results of the morphemic coding for each word revealed high inter-coder agreement. There was an 82% agreement across the three coders for W1s, and 73% for W2s. Pairwise correlations between individual coders were even higher, especially for W1. Coders A and B: 82% for W1 and 73% for W2; Coders A and C: 91% for W1 and 88% for W2; Coders B and C: 92% for W1 and 85% for W2. Mismatches between the coders' judgments were due to hesitations about the meaning of certain words or differences in the role attributed to etymology.

Our analyses focussed on the word pairs whose members were both judged as mono-morphemic ($N = 29$, e.g. *été-éther*) or bimorphemic ($N = 55$, e.g. *prêteuse-prêteur*) by at least two coders out of three. The number of mono-morphemic and bimorphemic pairs for each type of word pair is given in Table V.

— Insert Table V about here —

The assumption that vowel harmony is restricted to bimorphemic words leads us to predict the following patterns of variation. Using again ΔF_1 and ΔF_2 as acoustic measures of VH, we may expect ΔF_1, on the one hand, to have a larger *negative* value for bimorphemic than for monomorphemic pairs, that is to show a negative correlation with the number of morphemes. ΔF_2, on the other hand, should display a *positive* correlation with the number of morphemes for word pairs with a front V1, and a negative one for word pairs with a back V1. An exception to the latter case is raised by the *notice-nota* word pairs, for which the correlation is again expected to be positive. These predictions were assessed using the point-biserial correlation method, a type of bivariate correlation used when one of the two variables is dichotomous, i.e. categorical with only two possible categories[10], as was the case for the number of morphemes in our analyses. Because the unambiguously mono-morphemic and bi-morphemic word pairs were relatively scarce in our material, the point-biserial correlation coefficient (r_{pb}) between the number of morphemes and ΔF_1 or ΔF_2 was computed across all the available pairs, for each speaker and each V1 category. When ΔF_2 was the dependent variable, however, we left out the *notice-nota* word-pair set, since ΔF_2 showed for this set a pattern of variation opposite to that of the four other sets associated

19

with the back V1. The results are shown in Table VI.

— Insert Table VI about here —

Table VI shows that most of the correlation coefficients were statistically non-significant. There were only two exceptions to this general trend (Speakers S2 and S3, back V1, ΔF_2, significantly negative correlations consistent with the above predictions). This suggests that vowel harmony applies regardless of the presence or absence of a morpheme boundary between the two vowels, in accordance with Tranel (1987) and contrary to Dell (1973). However, it should be recalled that ΔF_1 and ΔF_2 were found to significantly vary in V1 as a function of V2 for certain V1-V2 combinations only (Tables III and IV). Since all the word pairs associated with each V1 were pooled together in the correlation analyses, it may be the case that these data contained too much statistical noise for the influence of morphological structure on vowel harmony to show up. Further comments on these issues are offered in the general discussion.

3.4 Potential relationship between lexical frequency on vowel harmony

In this section, we examine the potential influence of lexical frequency on the size of vowel harmony effects. Previous studies have shown that lexical frequency has an effect on lexical (Fidelholz, 1975; Hooper, 1976) and phonetic (Wright, 2003) vowel reduction and overall word durational shortening in English (see Jurafsky, 2003, for a review). For example, Wright (2003) found that vowels showed a greater degree of reduction in so-called easy CVC English words, with a high frequency relative to that of their phonological neighbors and a sparse neighborhood, than in hard words, with a low frequency relative to that of their phonological neighbors and a dense neighborhood. This phenomenon was attributed to the fact that easy words have less competitors than hard words in lexical access and are therefore intrinsically more intelligible. It may be assumed, as in Lindblom's (1990) hyper/hypo-articulation theory for example, that words that are easier to recognize by listeners will tend to undergo vowel reduction and overall shortening to a greater extent. From a historical point of view, Bybee (2001) suggests that sound change resulting from phonetic reduction does not simultaneously apply to every word in the lexicon, and

20

that high-frequency words are affected earlier than low and mid frequency words. The potential impact that these perceptual and historical factors may have on vowel harmony in French needs to be assessed.

In a classical, segmental approach to speech perception, vowel harmony may be regarded as causing V1 to deviate from a non-assimilated, abstract canonical form, hence making this vowel more difficult to recognize by listeners, which would lead to the prediction that high-frequency words will show a greater amount of vowel harmony than low-frequency words. By contrast, it may be assumed that vowel harmony results in the first vowel containing perceptually relevant cues to the second vowel, and that this will enhance the word's intelligibility. In a companion perceptual study (Nguyen, Fagyal & Cole, 2004) using part of the material described in the present paper as stimuli, we found that vowel harmony indeed facilitates the identification of V2, in certain circumstances at least, and may therefore speed up the recognition of the carrier word. On this account, one would expect vowel harmony to be more extensive in *low*-frequency, hard-to-recognize words, than in high-frequency words. Likewise, from a historical perspective, different proposals can be made with respect to the potential interplay between lexical frequency and vowel harmony. According to Bybee (2001), high frequency of usage encourages phonetic changes that involve the retiming of articulatory gestures, as this occurs in assimilation, in particular (Browman and Goldstein, 1986). If vowel harmony in French forms a standard case of assimilation, we might expect infrequent words to display less vowel harmony, as they would tend to remain "entrenched" in conservative patterns of pronunciation, whereas frequent words would show stronger assimilatory tendencies, as they tend to be more readily drawn into innovative phonetic changes. Note, however, that the ongoing trend in French could be towards the reduction of vowel harmony. The patterns we found in the present study were indeed more subtle than those described in earlier work (e.g. Dell, 1973; Fouché, 1959; Grammont, 1939; Malmberg, 1969). In addition, to what extent these patterns correspond to assimilation as classically defined (see below, Sections 3.5 and 4) is open to debate. Examining the relationships between lexical frequency and vowel harmony may therefore provide insight into both the perceptual relevance of VH, and the historical trend in which VH seems to

21

be engaged.

Lexical frequency estimates were collected for all words in our material (see Note 4 for details). The distribution of values was highly skewed towards the low-frequency range, with a median value of 0.45 occurrences per million words, and a maximum value of 214.51. Nineteen words had a frequency equal to 0^{11}. Average lexical frequency values were then computed for each word pair, and compared to the corresponding values of ΔF_1 and ΔF_2, using the Kendall rank-order correlation coefficient tau[12]. For each speaker and each V1 category, a one-tailed test was carried out across all the available pairs. When ΔF_2 was the predicted variable, however, the *notice-nota* word-pair set was left aside, for reasons already stated above. The results are shown in Table VII.

— Insert Table VII about here —

Table VII indicates that, in general, the correlation between lexical frequency and our acoustic measures of vowel harmony was very low. The $p < .05$ level of significance was reached in three cases only (Subject S3, back V1, ΔF_2; Subject S6, back V1, ΔF_1 and ΔF_2). The first of these was consistent with low-frequency words showing more vowel harmony than high-frequency words, whereas the two other cases displayed the opposite tendency (less vowel harmony for low- than for high-frequency words).

Two factors may have prevented a clearer pattern from emerging from our data. First, the concentration of the words within a restricted range in the bottom part of the lexical frequency scale probably made the relationships that this parameter may show with vowel harmony more difficult to capture. Second, because our measures of vowel harmony were associated with word pairs, as opposed to individual words, they were compared with the mean lexical frequency for the two members in each pair, which was taken as being representative of the lexical frequencies values associated with each of the two words. A comparison between the lexical frequencies for the first and second word across all the word pairs showed that this was not the case, however, as the correlation between the two series of values was close to 0 (0.02). Thus, the potential influence of lexical frequency on vowel harmony may need to be further explored using a more balanced set of words with respect to lexical frequency.

3.5 Further analyses: categoriality and gradience in vowel harmony

Symbolic phonological approaches to vowel harmony in French (Angoujard, 2004, 2006; Dell, 1973; Durand and Lyche, 2004; Selkirk, 1972) characterize vowel harmony as a phenomenon that affects the harmonizing vowel in an all-or-none fashion. Although VH is usually said to be optional and to occur more frequently in more informal speech styles, when it does occur, it is represented as being associated with a categorial difference in the height of the harmonizing vowel. In addition, however, it may be the case that more gradual variations are shown by the harmonizing vowel as a function of the detailed articulatory characteristics of the triggering vowel. In the word *été* [ete] for example, fine-grained differences in how the final vowel is phonetically realized may be reflected in the first vowel. Standard phonetic theory would refer to this as V-to-V coarticulation, as opposed to phonological assimilation, although to what extent this distinction applies to vowel harmony in French will be questioned in the general discussion. In the present section, we try to determine the relative contribution of gradient and categorial effects to vowel harmony.

The method we used consisted in attempting to predict the formant frequency values for V1 from both the phonological height and the formant frequency values of V2. Phonological height was a discrete parameter with two possible values ($\pm$low), whereas formant frequency was of course a continuous parameter. In these analyses, we only considered F_1 frequency for both V1 and V2, because of the relationship that may be established between F_1 and vowel height. The question we asked was how much of the variance in F_1 frequency in V1 could be accounted for by V2's phonological height, on the one hand, and by F_1 frequency in V2, on the other hand (see Jurafsky et al., 2001, for a similar approach to the study of function word duration in English).

For each speaker, and each word-pair set, a linear mixed-effects model (Baayen, forthcoming; Johnson, 2007; Pinheiro and Bates, 2000; Quené and van den Bergh, 2004) was designed with F_1 frequency in V1 as the predicted variable, the V2 category (non-low in the first member of each word pair vs low in the second member) and the frequency of F_1 in V2 as the two predictors (fixed effects), and one blocking factor (random effects), the word pair. Note that the repeated-measure ANOVAs presented in Section 3.2 can be seen as a special case of these more general models,

23

since the ANOVAs were equivalent to linear models containing only the V2 category as predictor. Table VIII contains the t and p values associated with each predictor in each model[13]. To take an example, Table VIII indicates that for Speaker S1 and for the front V1 before e/ε, F_1 frequency in V1 significantly varies depending on V2's phonological height ($t(95) = 4.79, p < 0.001$, see top leftmost cell).

— Insert Table VIII about here —

In a number of cases (e.g. Speaker S1, front V1 prior to i/a), neither of the two predictors was found to be significant, in spite of the V2 category having a significant effect in the corresponding ANOVA. This is attributable to the fact that the part of acoustic variance in V1 explained by the model is split between the predictors, and when this part is relatively small and/or shared in equal proportions by the predictors, this may prevent either of the predictors from reaching significance. (In such cases, and when the predictors were introduced into the model in a stepwise fashion, we found that the first one was indeed significant.) When either of the two predictors was significant, the relationship between this predictor and F_1 in V1 was always positive (i.e. pointing to a harmony between V1 and V2). V2 height was the only significant predictor in two cases, whereas F_1 in V2 was the only significant predictor in eleven cases, both predictors being significant in a further eight cases. Overall, therefore, gradient effects were either combined with or were stronger than categorial effects. In other words, when only one predictor had a significant effect on F_1 in V1, it was more likely to be F_1 frequency in V2 than V2's phonological height.

Patterns that emerged from these analyses are illustrated in Figure 8. In this figure, the relationship between F_1 in V2 and V1 is depicted separately for words containing a low V2, on the one hand, and a non-low V2, on the other hand. In the left panel, both F_1 frequency in V2 and V2's phonological height provided a significant contribution to accounting for the variance shown by F_1 frequency in V1. The two other panels correspond to situations where only one regressor was found to be significant, either V2's phonological height (middle panel) or F_1 frequency in V2 (right panel).

— Insert Figure 8 about here —

The left panel illustrates how F_1 frequency in V1 varied as a function of F_1 frequency in V2 in the *poseuse-poseur* word-pair set for Speaker S1. The first member of each word pair, which has a non-low V2 (/ø/, as in *poseuse*), is represented by a filled circle, and the second member, with a low V2 (/œ/, as in *poseur*), by an open circle. There was a 205 Hz decrease in the F_1 frequency of the non-low V2, relative to its low counterpart, which was paralleled by a 11 Hz decrease in F_1 frequency for V1, as reported in Table III. In addition, the slope of the local regression line was significantly positive (0.56, $t = 3.81$, df = 56, $p < 0.001$) for the words containing a non-low V2. Figure 8 reveals further interesting features in our data. In the middle panel, the slope of the local regression line was significantly negative for the low V2 (-0.45, $t = -2.75$, df = 17, $p < 0.05$). Why the slope was negative in this particular case is unclear and may be due to word-specific characteristics in the phonetic realization of the *dévot-dévote* word set (notice that this set contained six word pairs only). However, this pattern may explain why V2's phonological height was the only significant predictor of F_1 frequency in V1 in this word set, as F_1 in V2 was correlated with F_1 in V1 in opposite ways depending on whether V2 was non-low or low. A negative, albeit non-significant regression slope (-0.57, $t = -1.83$, df = 17, $p = 0.08$) was also found for the low V2 for Speaker S5 in the *dévot-dévote* set, which constituted the only other case in which V2's phonological height was the sole significant predictor.

It is also evident from the left and right panels in Figure 8 that the slope of the regression line is significantly positive for the non-low V2 only. This reflects a general trend in our data: in most cases where F_1 in V2 was found to explain a significant proportion of the variance for F_1 in V1, this contribution proved stronger for the non-low V2. There were in fact only three exceptions to this, i.e. Speaker S2, front V1 before e/ɛ and back V1 before o/ɔ (significantly positive regression slope for the low V2) and Speaker S5, back V1 before e/ɛ (significantly positive regression slopes for both the non-low and the low V2). The closer relationship between F_1 in V1 and V2, when V2 is a non-low vowel as opposed to a low one, may be explained by the fact that small variations in the vocal-tract cross-sectional area in the vicinity of the constriction for the vowel may have a

25

stronger impact on formant frequencies when the tongue dorsum is closer to the roof of the oral cavity. Regardless of the appropriate articulatory interpretation, the above finding provides a clear indication that the acoustic characteristics of V1 were influenced by fine-grained, infra-phonemic variations in how V2 was realized.

To summarize, a positive relationship was generally found between F_1 frequency in V2 vs V1, particularly when V2 was a non-low vowel. In a number of cases, V2's phonological height also explained part of the variance in F_1 frequency for V1. While further analyses may be warranted, this pattern suggests that gradient effects are either combined with or tend to override categorial effects in vowel harmony in French.

4   General discussion

In this paper, we examined the acoustic aspects of vowel harmony, understood as regressive vowel-to-vowel assimilation in the two vowels of disyllabic words, in six speakers representing two regional varieties of French, Northern and Southern French spoken in France. Our study focused on 107 disyllabic word pairs in which the word-initial vowel (V1) was phonemically either /e/ or /o/, and the word-final stressed vowel (V2) alternated between /e-ɛ/, /ø-œ/, /o-ɔ/ or /i-a/. Our main results can be summarized as follows. In general, V1 was higher when followed by a high-mid vowel than by a low-mid vowel. In addition, V1 was both higher (lower F_1 frequency) and more fronted (higher F_2 frequency) before /i/ than before /a/. These variations affected both the front and back vowels in the first syllable, and were independent of the presence or absence of a morphemic boundary between the two syllables and the carrier words' frequencies. They were more systematic in some speakers, particularly those speaking a Northern French variety. The data also indicated that gradient assimilatory effects predominated upon categorial harmony effects.

Vowel harmony has often been portrayed as a tendency towards closure that affects the front unrounded mid vowel /ɛ/ prior to the front non-low vowels /i, y, e/. Contrary to this assumption, our data show that vowel harmony can also involve back mid vowels, both as targets (e.g. *potée-poterne* word-pair set) and triggers (e.g. *dévot-dévote*). Whether vowel harmony can only take the

form of a closing gesture (before a high-mid or high vowel) is difficult to establish from our data. The closing-gesture-only hypothesis seems to be based on impressionistic three-way comparisons between a monosyllabic base form (e.g. *fête* /fɛt/ "fair") and two disyllabic derived or inflected forms with a suffix containing a low vowel (e.g. *fêtard* "reveller"), on the one hand, and a non-low vowel (e.g. *fêter* "to celebrate"), on the other hand. It is assumed that the non-low vowel in *fêter*, to take this example, has an assimilatory influence on the preceding vowel (/fete/) while the low vowel in *fêtard* would allow the quality of the vowel in the base to be preserved (/fɛtaʁ/). This hypothesis could not be directly tested here since our material was not designed to compare base forms to derived or inflected forms. Note, however, that there was an overall tendency for V1 to be closer to a mid-high vowel than to a mid-low vowel, regardless of the acoustic properties of the following vowel. As shown by Figures 2-7, the front and back V1 were in most cases located in the vicinity of the points associated with /e/ and /o/ respectively on our reference grid. More material would be needed to determine whether the height of low vowels in base forms is preserved by our speakers in suffixed forms, in spite of this general tendency.

Using the Optimality Theory framework, Durand and Lyche (2004) propose that the vowel systems in different varieties of French are governed by an interplay between three main constraints: the *loi de position*, vowel harmony, and faithfulness to the root. Durand and Lyche suggest that in the two local varieties examined in their study (coastal Southern French and French spoken in Grenoble), the *loi de position* is ranked at the top of the constraint hierarchy and faithfulness to the root at the bottom, with vowel harmony having an intermediate ranking. Our data are partially consistent with this hierarchy for both the Southern and Northern speakers. Since the first syllable was open in all of target words in our corpus, and given that mid vowels are expected to be high-mid in open syllables according to the *loi de position*, the above trend for V1 to be relatively high overall could be taken as a reflection of the *loi de position*. Under this assumption, the relatively small size of the observed vowel harmony effects could be interpreted as a conflict between vowel harmony and the higher-ranked *loi de position*. It is difficult to assess empirically the potential impact of a constraint such as faithfulness to the root on the vowels' surface forms, as

27

pointed out by Durand and Lyche (2004) themselves. This is due to the fact that the corresponding underlying forms may be unspecified for height or tenseness. For part of our material, underlying forms can be thought of as closed syllables with a low-mid vowel, e.g. *flèche* /flɛʃ/ "arrow", whose suffixed forms are the word pair *fléché-fléchette*. However, the relative proportion of such suffixed vs non-suffixed forms in each word-pair set was not well balanced, and a larger corpus would again be necessary to determine whether faithfulness to the root can hinder vowel harmony (see Méjean, 2007, for a pilot study in that direction).

As already indicated, preliminary evidence suggests that the vowel harmony patterns described in the present study are perceptually salient. In Nguyen, Fagyal & Cole (2004), six trained phoneticians were presented with initial syllables extracted from the *épice-épate* and *notice-nota* word sets, as spoken by our prototypical Parisian speaker (S2), in a modified version of the ABX task (Majors, 1998). These syllables were judged to sound more like each other when originating from two repetitions of the same disyllabic word (e.g. *no(tice)*, rep. 1, compared with *no(tice)*, rep. 2) relative to a pair of words with different final syllables (e.g. *no(tice)* compared with *no(ta)*). Since the most systematic difference between the two members of each word pair lay in the final vowel (/i/ for one member and /a/ for the other), it may be assumed that the observed response patterns reflected the effect that final vowels have on how initial vowels (and possibly word-onset consonant(s), if any), are produced. This effect appeared to be more perceptually salient than the acoustic variations that occur between repetitions for each word. Another experiment aimed to examine whether vowel harmony facilitates the identification of the final vowel in a disyllabic word. Drawing on previous studies carried out by Martin and Bunnell (1981, 1982) on the perception of V-to-V anticipatory coarticulation in English, Nguyen et al. (2004) had 16 naive listeners perform a V2 detection task in words whose initial syllable either was the original one, or was cross-spliced from another word with a different final syllable. When the first vowel contained conflicting cues to the second vowel, the subjects' responses were delayed relative to the control condition, for part of the word pairs. These results provide partial support for the assumption that listeners are sensitive to regressive vowel harmony effects and use them to predict the identity of an upcoming vowel.

It may in fact be hypothesized that vowel harmony in French is, at least in part, perceptually motivated, i.e. that speakers produce it with a view to making the identification of the carrier word easier for the listeners. Because the contrast between words containing a nonlow vs low V2, is realized as a set of acoustic properties which extend to the preceding vowel at least (e.g., for both V1 and V2, a more closed vowel quality when V2 is high-mid, and a more open vowel quality when V2 was low-mid), vowel harmony may help speed up the recognition of the carrier words and make it more resistant to noise. Perceptual factors have been invoked to explain the emergence of phonological vowel harmony as a way to enhance the perceptual salience of the harmonic feature (Kaun, 1995).

The present study focused on the acoustics of vowel harmony, and the articulatory characteristics that are associated with it remain to be determined. Few empirical data are available that would allow us to shed more light on these characteristics. These data include classic work by Straka (1965) and Bothorel et al. (1986), together with the insightful comments offered by Coveney (2001). On the whole, they are consistent with the idea that vowel harmony primarily entails variations in tongue height, along with related displacements of the tongue position along the front-back axis. To the extent that F_1 and F_2 are correlated with tongue height and front-back position, respectively, our acoustic measures also lend support to this assumption. Articulatory investigations would of course be required to further our understanding of the dynamics of the tongue and jaw movements in vowel harmony. Another issue that was not dealt with in the present work and that articulatory measurements would allow to address directly concerns the potential blocking effect that intervocalic consonants, particularly dorsal ones, may have on vowel harmony.

Our data also raise questions regarding symbolic phonological approaches to vowel harmony in French. The effects we observed are subtle and gradual, and they do not correspond to the traditional characterization of vowel harmony as a plus-minus switch in the height of the harmonizing vowel. Infra-phonemic variations in F_1 for V2 tended to absorb most of the variance in V1's own F_1 frequency, in conjunction with or at the expense of V2's phonological height. Because the linear mixed-effects models allowed V2's phonological height and F_1 frequency to be simultaneously

29

fitted against F_1 frequency in V1, these models provide, in our view, powerful evidence that vowel harmony in French is a primarily gradient, as opposed to categorial, phenomenon. However, vowel harmony patterns in our data also showed important differences with respect to V-to-V coarticulation, as it is portrayed in standard phonetic theory at least. These patterns were to some extent speaker-dependent, and appeared more systematic in Northern than Southern French. In addition, different patterns of variation in V1 were associated with the mid-vowel constrasts and with the i/a contrast in V2 position. Importantly, our acoustic data suggest that the back V1 was *more retracted* before a high-mid front than before a low-mid front vowel (e.g. *potée* vs *poterne*), while being *more fronted* before the high front vowel /i/ than before the low vowel /a/ (e.g. *notice* vs *nota*). Non-final mid vowels therefore behave prior to another mid vowel in a rather specific way, which does not seem to conform to general, language-independent tendencies in V-to-V coarticulation.

Thus, in spite of its being a fine-grained, gradient phenomenon, it may well be that, at least in certain varieties of French, vowel harmony is part of the speakers' phonological knowledge. One way to get around what may appear, in a symbolic phonological framework, to be a paradox, is to assume that lexical representations include continuous, detailed phonetic information associated with individual words, as is the case in usage-based phonology (Bybee, 2001; Pierrehumbert, 2002) for example. Articulatory phonology (Browman and Goldstein, 1992; Goldstein and Fowler, 2003) also constitutes a promising avenue for modelling vowel harmony in French, since gradience is integral to the gestural scores of which lexical representations are made up in this approach, and also because recent developments combining articulatory phonology with the theory of non-linear dynamical systems have endeavored to integrate both continuous phonetic and discrete phonological aspects of vowel harmony in various languages (Benus, 2005; Gafos, 2006). Thus, the present work may have interesting general implications for how the relationship between phonology and phonetics is defined. In a standard view, phonology is associated with discrete and abstract categories, while phonetics refers to inherently continuous and gradient patterns. As shown by Cohn (2005), however, there is an alternative view, which conceives phonetics and phonology

30

as taking place on a continuum ranging from more fine-grained to more granular dimensions of speech. Vowel harmony in French may provide support for this view, as an instance of fine phonetic variation that is incorporated into the speaker's mental lexicon.

As for the question whether VH, understood as regressive vowel-to-vowel assimilation, is manifest in both Northern and Southern French, our answer should be: yes. All speakers in our corpus showed at least some tendency towards VH, but the degree to which they exhibited such a tendency varied considerably. At the strongest endpoint of the continuum we find a prototypical Parisian speaker who can be opposed to a prototypical Southern speaker, showing the least vowel-to-vowel assimilation. Notwithstanding individual speaker-dependent variations, this finding is suggestive of the fact that VH is present in both speakers' phonological representations but stronger for Northern than for Southern speakers. Thus, rather than yielding categorical distinctions, VH in these two varieties of French corresponds to lesser vs stronger gradient patterns of assimilation. The tendency of VH to be stronger in Northern French speakers is further supported by the observation that the two Southern speakers who exhibited VH were those who seemed to style shift during the lab recording session: they adopted a formal reading style that, based on auditory observation at least, shared several features with Standard French that is phonologically closer to Northern French varieties (see Section 2.2). To come back to Ohala's far-reaching hypothesis, only time can tell whether the Northern dialect, exhibiting the strongest assimilatory pattern at this point in time, will evolve towards more categorical VH-like assimilatory distinctions in the future.

Acknowledgements

Notes

[1] Since investigations have focussed mainly on the effect of the word-final, primary stressed vowel on the preceding vowel, comparatively little attention has been paid to whether VH can extend over a larger domain in polysyllabic words, although Grammont (1939) implicitly suggested that this might indeed be the case (e.g. *bégaiement* [bɛgɛmɑ̃] "stuttering" vs. *bégayer* [begeje] "to stutter").

[2] It is noteworthy that informal observations by native speakers seem to contradict this observation. The first vowel in the first name *Kevin*, for instance, can give the impression of fronting and closing of the /e/ to /i/ in the speech of native speakers of Parisian French. The open vowel /a/ in word-initial position seems to undergo fronting and closing, as well (J. Vaissière, personal communication).

[3] "General French [which] can be considered equivalent to Received Pronunciation for British English or General American for American English. General French is characterized by the absence of dialectal marks and is mainly used by educated people as well as by professional radio and television [announcers]" (Di Cristo, 1998, 195).

[4] A complete list of the words contained in the corpus is available at `www.lpl.univ-aix.fr/~nguyen/v_harm.html`.

[5] This division reflects the well-known, and perceptually clear, split between "oïl" and "oc" dialect areas in France (see e.g. Lodge, 1993). It does not control for fine-grade differences within these regional varieties.

[6] These errors were mostly due to the proximity of F_1 and F_2 in /o/, while the short duration of the vowel in the first syllable may have also contributed to making the two formants more difficult to identify for the formant tracker.

[7] We considered using other auditory scales such as the mel scale or the more recently developed ERB scale (Moore and Glasberg, 1983), which are sometimes used rather than the Bark scale (see for example Rosner and Pickering, 1994). However, the results of the ANOVAs we performed on the data were essentially the same regardless of which auditory scale was used. That is, substituting mels or ERBs to Barks made little difference to the statistical significance of the effects that the independent variables were found to have on the dependent variables. Barks were used in the present paper to facilitate comparisons with previous studies on the acoustic and auditory characteristics of vowels in various languages (e.g. Ladefoged and Maddieson, 1990; Stevens, 1998; Syrdal and Gopal, 1986) which also employed that scale.

[8] Previous work on the perception of vowel height (e.g. Hoemeke and Diehl, 1994; Syrdal and Gopal, 1986; Traunmüller, 1981) has suggested that $[z(F_1) - z(f_0)]$ is a more reliable cue to this contrast than $z(F_1)$ alone both across and within speakers. More specifically, $[z(F_1) - z(f_0)]$ may contribute to factor out speaker-dependent variations in the frequency of F_1, insofar as these are correlated with variations in the fundamental frequency. In the present study, however, statistical analyses were performed for each speaker separately, which made such an inter-speaker normalization procedure unnecessary. It has also been assumed that $[z(F_1) - z(f_0)]$ varies to a greater extent than $z(F_1)$ depending on vowel height because of the well-known correlation shown by f_0 itself with vowel height in a given speaker (for French, see Di Cristo, 1980). However, differences in intrinsic fundamental frequency

33

are expected to be minimal among the mid vowels on which this study is focused. Thus, looking at the mid vowels contained at the end of the target words in our material, we found that whereas f_0 varies in the expected direction in most cases (i.e. it was higher when the vowel was mid-high rather than mid-low, all other things being equal), the mean difference in f_0 between mid-high and mid-low vowels, all speakers taken together, was +2.5 Hz only. By contrast, the mean difference in F_1 between mid-high and mid-low vowels was –143 Hz. We concluded that f_0 was not likely to provide a significant contribution to differentiating mid vowels along the height dimension.

[9]Phonetic symbols for these vowels are not displayed due to lack of space but can be found in Figure 1.

[10]A point-biserial correlation is a Pearson correlation when the dichotomous variable is coded with 0 for one category and 1 for the other, see e.g. Field (2003) for further detail.

[11]This means that they did not occur in the textual database lexical frequencies were extracted from.

[12]The Kendall tau non-parametric correlation test was preferred over other non-parametric tests (e.g. Spearman's), because it is less permissive with small data sets that contain a relatively high number of tied ranks (scores that, once ranked, end up with the same rank).

[13]The full set of results for the linear mixed-effects models is available at `www.lpl.univ-aix.fr/~nguyen/v_harm.html`.

References

Angoujard, J.-P. (2004). [DOM] comme domaine. In *Proceedings of the IVth Linguistic Studies Workshop*, pages 157–162, Nantes, France.

Angoujard, J.-P. (2006). *Phonologie déclarative*. CNRS Éditions, Paris.

Baayen, H. (forthcoming). *Analyzing Linguistic Data: A Practical Introduction to Statistics*. Cambridge University Press, Cambridge, UK. Preliminary version available at www.mpi.nl/world/persons/private/baayen/publications/baayenCUPstats.pdf.

Benus, S. (2005). *Dynamics and transparency in vowel harmony*. PhD thesis, New York University, New York.

Bothorel, A., Simon, P., Wioland, F., and Zerling, J.-P. (1986). *Cinéradiographie des voyelles et consonnes du français*. Travaux de l'Institut de Phonétique de Strasbourg, Strasbourg.

Boula de Mareüil, P. and Fagyal, Z. (2000). Autour de l'harmonie vocalique en français. In *Actes des XXIIIèmes Journées d'Étude sur la Parole*, pages 85–88, Aussois.

Browman, C. and Goldstein, L. (1986). Towards an articulatory phonology. *Phonology Yearbook*, 3:219–252.

Browman, C. and Goldstein, L. (1992). Articulatory phonology: an overview. *Phonetica*, 49:155–180.

Bybee, J. (2001). *Phonology and Language Use*. Cambridge University Press, Cambridge, UK.

Casagrande, J. (1984). *The Sound System of French*. Georgetown Univ. Press, Washington, D.C.

Cohn, A. (2005). Gradience and categoriality in sound patterns. Paper presented at the Workshop on Phonological Systems and Complex Adaptive Systems, Lyons, France, 4–6 July 2005.

Coveney, A. (2001). *The Sounds of Contemporary French: Articulation and Diversity*. Elm Bank Publications, Exeter, UK.

Dell, F. (1973). *Les règles et les sons.* Hermann, Paris.

Di Cristo, A. (1980). La fréquence fondamentale intrinsèque des voyelles du français. *Travaux de l'Institut de Phonétique d'Aix-en-Provence*, 7:183–207.

Di Cristo, A. (1998). Intonation in French. In Hirst, D. and Di Cristo, A., editors, *Intonation Systems. A Survey of Twenty Languages*, pages 195–218. Cambridge University Press, Cambridge, UK.

Durand, J. (1990). *Generative and Non-Linear Phonology.* Longman, London, UK.

Durand, J. and Lyche, C. (2004). Structure et variation dans quelques systèmes vocaliques du français: l'enquête Phonologie du français contemporain (PFC). In Coveney, A. and Sanders, C., editors, *Variation et francophonie*, pages 217–240. L'Harmattan, Paris.

Fagyal, Z., Nguyen, N., and Boula de Mareüil, P. (2003). From dilation to coarticulation: is there vowel harmony in French? *Studies in Linguistic Sciences*, 32:1–21.

Fidelholz, J. (1975). Word frequency and vowel reduction in English. *CLS-75*, pages 200–213. University of Chicago.

Field, A. (2003). *Discovering Statistics Using SPSS for Windows.* SAGE Publications, London.

Fouché, P. (1959). *Traité de prononciation française.* Klincksieck, Paris. 2nd edition.

Gafos, A. (2006). Dynamics in grammar: Comment on Ladd and Ernestus & Baayen. In Goldstein, L., Whalen, D., and Best, C., editors, *Papers in Laboratory Phonology 8: Varieties of Phonological Competence*, pages 51–79. Mouton de Gruyter, Berlin, Germany.

Goldstein, L. and Fowler, C. (2003). Articulatory phonology: a phonology for public language use. In Meyer, A. and Schiller, N., editors, *Phonetics and Phonology in Language Comprehension and Production: Differences and Similarities*, pages 159–207. Mouton de Gruyter, Berlin, Germany.

Grammont, M. (1914). *La prononciation française: traité pratique.* Delagrave, Paris.

Grammont, M. (1939). *Traité de phonétique*. Delagrave, Paris. 2nd edition.

Hoemeke, K. and Diehl, R. (1994). Perception of vowel height: The role of $F_1 - F_0$ distance. *Journal of the Acoustical Society of America*, 96:661–674.

Hooper, J. (1976). Word frequency in lexical diffusion and the source of morphophonological change. In Christie, W., editor, *Current Progress in Historical Linguistics*, pages 96–105. North Holland, Amsterdam.

Hyman, L. (2002). Is there a right-to-left bias in vowel harmony? In *9th International Phonology Meeting*, Vienna, Austria.

Johnson, K. (2007). *Quantitative Methods in Linguistics*. in preparation; preliminary version on-line available at `linguistics.berkeley.edu/~kjohnson/quantitative/`.

Jurafsky, D. (2003). Probabilistic modeling in psycholinguistics: linguistic comprehension and production. In Bod, R., Hay, J., and Jannedy, S., editors, *Probabilistic Linguistics*, pages 39–95. MIT Press, Cambridge, Mass.

Jurafsky, D., Bell, A., Gregory, M., and Raymond, W. (2001). Probabilistic relations between words: Evidence from reduction in lexical production. In Bybee, J. and Hopper, P., editors, *Frequency and the Emergence of Linguistic Structure*, pages 229–254. John Benjamins, Amsterdam.

Kaun, A. (1995). *The Typology of Rounding Harmony: An Optimality-Theoretic Approach*. PhD thesis, UCLA.

Klatt, D. and Klatt, L. (1990). Analysis, synthesis, and perception of voice quality variations among female and male talkers. *Journal of the Acoustical Society of America*, 87:820–857.

Ladefoged, P. and Maddieson, I. (1990). Vowels of the world's languages. *Journal of Phonetics*, 18:93–122.

Landick, M. (1995). The mid-vowels in figures: hard facts. *The French Review*, 69:88–102.

Landick, M. (2004). *Enquête sur la prononciation du français de référence. Les voyelles moyennes et l'harmonie vocalique.* L'Harmattan, Paris.

Lindblom, B. (1990). Explaining phonetic variation: a sketch of the H&H theory. In Hardcastle, W. and Marchal, A., editors, *Speech Production and Speech Modelling*, pages 403–439. Kluwer, Dordrecht.

Lodge, A. (1993). *French: from Dialect to Standard.* Routledge, London.

Majors, T. (1998). *Stress dependent harmony: Phonetic origin and phonological analysis.* PhD thesis, University of Texas, Austin, TX.

Malderez, I. (1995). *Contribution à la synchronie dynamique du français contemporain: le cas des voyelles orales arrondies.* PhD thesis, Université de Paris 7.

Malmberg, B. (1969). *Phonétique française.* Hermods, Malmö.

Malmberg, B. (1974). *Manuel de phonétique générale.* Picard, Paris.

Martin, J. and Bunnell, H. (1981). Perception of anticipatory coarticulation effects. *Journal of the Acoustical Society of America*, 69:559–567.

Martin, J. and Bunnell, H. (1982). Perception of anticipatory coarticulation effects in vowel-stop consonant-vowel sequences. *Journal of Experimental Psychology: Human Perception and Performance*, 8:473–488.

Martinet, A. (1945). *La prononciation du français contemporain.* Droz, Paris.

Mettas, O. (1979). *La prononciation parisienne: aspects phoniques d'un sociolecte parisien (du Faubourg Saint-Germain à la Muette).* Éditions du CNRS, Paris.

Méjean, L. (2007). L'harmonie vocalique en français standard. Master's thesis, Aix-Marseille Université, Aix-en-Provence.

Moore, B. and Glasberg, B. (1983). Suggested formulae for calculating auditory-filter bandwidths and excitation patterns. *Journal of the Acoustical Society of America*, 74:750–753.

Nguyen, N., Fagyal, Z., and Cole, J. (2004). Perceptual relevance of long-domain phonetic dependencies. In *Proceedings of the IVth Linguistic Studies Workshop*, pages 173–178, Nantes, France. (On-line available at http://www.lettres.univ-nantes.fr/lling/jel2004/papers/actes-jel2004.pdf).

Ohala, J. (1994). Towards a universal, phonetically-based theory of vowel harmony. In *Proceedings of the International Conference on Speech and Language Processing (ICSLP '94)*, pages 491–494, Yokohama, Japan.

Öhman, S. (1966). Coarticulation in VCV utterances: Spectrographic measurements. *Journal of the Acoustical Society of America*, 39:151–168.

Pierrehumbert, J. (2002). Word-specific phonetics. In Gussenhoven, C. and Warner, N., editors, *Papers in Laboratory Phonology VII*, pages 101–140. Mouton de Gruyter, Berlin, Germany.

Pinheiro, J. and Bates, D. (2000). *Mixed-Effects Models in S and S-Plus.* Springer-Verlag, New York.

Quené, H. and van den Bergh, H. (2004). On multi-level modeling of data from repeated measures designs: a tutorial. *Speech Communication*, 43:103–121.

Rosner, B. and Pickering, J. (1994). *Vowel Perception and Production.* Oxford University Press, Oxford, UK.

Selkirk, E. (1972). *The Phrase Phonology of English and French.* PhD thesis, MIT. Published by the Indiana University Linguistics Club, 1981.

Stevens, K. (1998). *Acoustic Phonetics.* MIT Press, Cambridge, Mass.

Straka, G. (1965). *Album phonétique.* Presses de l'Université Laval, Québec.

Syrdal, A. and Gopal, H. (1986). A perceptual model of vowel recognition based on the auditory representation of American English vowels. *Journal of the Acoustical Society of America*, 79:1086–1100.

Tranel, B. (1987). *The Sounds of French: An Introduction.* Cambridge Univ. Press, Cambridge, UK.

Traunmüller, H. (1981). Perceptual dimension of openness in vowels. *Journal of the Acoustical Society of America*, 69:1465–1475.

Traunmüller, H. (1990). Analytical expressions for the tonotopic sensory scale. *Journal of the Acoustical Society of America*, 88:97–100.

Walker, D. (2001). *French Sound Structure.* Univ. Calgary Press, Calgary.

Wright, R. (2003). Factors of lexical competition in vowel articulation. In Local, J., Ogden, R., and Temple, R., editors, *Papers in Laboratory Phonology, VI: Phonetic Interpretation*, pages 75–87. Cambridge University Press, Cambridge, UK.

Table I: Types and number of word pairs.

word-pair set	V1	V2	n
été /ete/ - *éther* /etɛʁ/	front	e/ɛ	14
prêteuse /pʁetøz/ - *prêteur* /pretœʁ/	front	ø/œ	8
dévot /devo/ - *dévote* /devɔt/	front	o/ɔ	6
potée /pote/ - *poterne* /potɛʁn/	back	e/ɛ	26
poseuse /pozøz/ - *poseur* /pozœʁ/	back	ø/œ	19
auto /oto/ - *automne* /otɔn/	back	o/ɔ	4
épice /epis/ - *épate* /epat/	front	i/a	14
notice /notis/ - *nota* /nota/	back	i/a	16

Table II: Gender, age and place of birth of the speakers

Speaker	Gender	Age (years)	Place of birth
S1	female	31	Lamballe (Brittany)
S2	male	32	Paris
S3	male	21	Paris
S4	female	34	Salon de Provence
S5	male	39	Salon de Provence
S6	female	31	Orange

Table III: Mean difference in F_1 and in F_2 for V1 depending on V2 for each speaker. ΔF_1 (resp. ΔF_2) is positive when F_1 (resp. F_2) is higher in V1 if V2 is mid-high or high as opposed to mid-low or low. Statistically significant differences ($p < 0.05$) are in bold. Frequency values given in Hz.

Speaker	V1	Formant	V2 e/ɛ	V2 ø/œ	V2 o/ɔ	V2 i/a
S1	e	ΔF_1	−13	−22	−5	**−32**
	e	ΔF_2	**+73**	**+85**	**+73**	**+117**
	o	ΔF_1	−12	**−11**	−12	−17
	o	ΔF_2	**−62**	**−56**	−64	**+36**
S2	e	ΔF_1	**−24**	**−31**	−28	**−37**
	e	ΔF_2	+30	+28	+10	**+53**
	o	ΔF_1	**−28**	**−16**	**−52**	**−35**
	o	ΔF_2	**−60**	−34	**−177**	**+63**
S3	e	ΔF_1	**−42**	**−55**	**−39**	**−50**
	e	ΔF_2	**+59**	**+56**	+22	**+63**
	o	ΔF_1	**−42**	**−41**	−31	**−43**
	o	ΔF_2	**−92**	−25	**−119**	**+98**
S4	e	ΔF_1	0	−7	0	**−12**
	e	ΔF_2	+36	−2	−74	+70
	o	ΔF_1	+2	**−10**	−4	−5
	o	ΔF_2	−19	−21	−105	**+30**
S5	e	ΔF_1	**−26**	−7	**−9**	**−32**
	e	ΔF_2	**+49**	−16	+21	**+96**
	o	ΔF_1	**−28**	−10	−40	**−40**
	o	ΔF_2	**−46**	−4	**−97**	**+46**
S6	e	ΔF_1	−11	+3	+6	**−24**
	e	ΔF_2	**+51**	+25	+7	**+51**
	o	ΔF_1	**−29**	+11	**−78**	−20
	o	ΔF_2	**−107**	+18	**−174**	−9

Table IV: F and p values associated with formant frequency differences in Table III.

Speaker	V1	Formant	V2			
			e/ɛ	ø/œ	o/ɔ	i/a
S1	e	ΔF_1	$F(1,13) = 2.295$, NS	$F(1,7) = 2.227$, NS	$F(1,5) = 1.165$, NS	$F(1,11) = 5.516$, $p < 0.05$
		ΔF_2	$F(1,13) = 32.812$, $p < 0.001$	$F(1,7) = 151.838$, $p < 0.001$	$F(1,5) = 24.343$, $p < 0.01$	$F(1,11) = 117.804$, $p < 0.001$
	o	ΔF_1	$F(1,25) = 3.492$, NS	$F(1,18) = 5.583$, $p < 0.05$	$F(1,3) = 0.491$, NS	$F(1,15) = 2.954$, NS
		ΔF_2	$F(1,25) = 41.49$, $p < 0.001$	$F(1,18) = 30.321$, $p < 0.001$	$F(1,3) = 7.354$, NS	$F(1,15) = 7.281$, $p < 0.05$
S2	e	ΔF_1	$F(1,13) = 10.814$, $p < 0.01$	$F(1,7) = 7.456$, $p < 0.05$	$F(1,5) = 5.369$, NS	$F(1,13) = 18.311$, $p < 0.01$
		ΔF_2	$F(1,13) = 3.76$, NS	$F(1,7) = 4.941$, NS	$F(1,5) = 0.37$, NS	$F(1,13) = 18.288$, $p < 0.01$
	o	ΔF_1	$F(1,25) = 15.479$, $p < 0.01$	$F(1,18) = 5.74$, $p < 0.05$	$F(1,3) = 10.634$, $p < 0.05$	$F(1,15) = 13.549$, $p < 0.01$
		ΔF_2	$F(1,25) = 31.23$, $p < 0.001$	$F(1,18) = 2.361$, NS	$F(1,3) = 31.369$, $p < 0.05$	$F(1,15) = 26.96$, $p < 0.001$
S3	e	ΔF_1	$F(1,13) = 28.906$, $p < 0.001$	$F(1,7) = 11.38$, $p < 0.05$	$F(1,5) = 16.733$, $p < 0.01$	$F(1,13) = 67.973$, $p < 0.001$
		ΔF_2	$F(1,13) = 24.949$, $p < 0.001$	$F(1,7) = 16.015$, $p < 0.01$	$F(1,5) = 3.538$, NS	$F(1,13) = 55.432$, $p < 0.001$
	o	ΔF_1	$F(1,25) = 71.974$, $p < 0.001$	$F(1,18) = 47.908$, $p < 0.001$	$F(1,3) = 2.346$, NS	$F(1,15) = 21.771$, $p < 0.001$
		ΔF_2	$F(1,25) = 50.437$, $p < 0.001$	$F(1,18) = 2.191$, NS	$F(1,3) = 10.983$, $p < 0.05$	$F(1,15) = 31.849$, $p < 0.001$
S4	e	ΔF_1	$F(1,13) = 0.001$, NS	$F(1,7) = 0.945$, NS	$F(1,5) = 0.008$, NS	$F(1,13) = 5.777$, $p < 0.05$
		ΔF_2	$F(1,13) = 2.147$, NS	$F(1,7) = 0.009$, NS	$F(1,5) = 5.077$, NS	$F(1,13) = 2.225$, NS
	o	ΔF_1	$F(1,25) = 0.262$, NS	$F(1,18) = 12.617$, $p < 0.01$	$F(1,3) = 4.895$, NS	$F(1,15) = 3.382$, NS
		ΔF_2	$F(1,25) = 2.203$, NS	$F(1,18) = 1.417$, NS	$F(1,3) = 9.954$, NS	$F(1,15) = 11.456$, $p < 0.01$
S5	e	ΔF_1	$F(1,13) = 51.543$, $p < 0.001$	$F(1,7) = 1.143$, NS	$F(1,5) = 6.97$, $p < 0.05$	$F(1,13) = 30.716$, $p < 0.001$
		ΔF_2	$F(1,13) = 24.118$, $p < 0.001$	$F(1,7) = 1.222$, NS	$F(1,5) = 1.62$, NS	$F(1,13) = 117.883$, $p < 0.001$
	o	ΔF_1	$F(1,25) = 76.074$, $p < 0.001$	$F(1,18) = 4.227$, NS	$F(1,3) = 4.203$, NS	$F(1,15) = 37.239$, $p < 0.001$
		ΔF_2	$F(1,25) = 20.449$, $p < 0.001$	$F(1,18) = 0.085$, NS	$F(1,3) = 77.491$, $p < 0.01$	$F(1,15) = 28.12$, $p < 0.001$
S6	e	ΔF_1	$F(1,13) = 1.093$, NS	$F(1,7) = 1.298$, NS	$F(1,5) = 0.171$, NS	$F(1,13) = 14.431$, $p < 0.01$
		ΔF_2	$F(1,13) = 9.777$, $p < 0.01$	$F(1,7) = 0.47$, NS	$F(1,5) = 0.115$, NS	$F(1,13) = 7.893$, $p < 0.05$
	o	ΔF_1	$F(1,25) = 12.805$, $p < 0.01$	$F(1,18) = 2.197$, NS	$F(1,3) = 20.278$, $p < 0.05$	$F(1,15) = 1.64$, NS
		ΔF_2	$F(1,25) = 31.296$, $p < 0.001$	$F(1,18) = 0.908$, NS	$F(1,3) = 98.154$, $p < 0.01$	$F(1,15) = 0.15$, NS

44

Table V: Number of mono-morphemic and bimorphemic pairs in each word-pair set.

word-pair set	V1	V2	1 morph.	2 morph.
été /ete/ - *éther* /etɛʁ/	front	e/ɛ	2	7
prêteuse /pʁetøz/ - *prêteur* /pʁetœʁ/	front	ø/œ	0	8
dévot /devo/ - *dévote* /devɔt/	front	o/ɔ	5	0
potée /pote/ - *poterne* /potɛʁn/	back	e/ɛ	5	15
poseuse /pozøz/ - *poseur* /pozœʁ/	back	ø/œ	0	19
auto /oto/ - *automne* /otɔn/	back	o/ɔ	2	0
épice /epis/ - *épate* /epat/	front	i/a	6	5
notice /notis/ - *nota* /nota/	back	i/a	9	1

Table VI: Relationship between the morphological make-up of the two members of each word pair and the size of vowel harmony effects for each speaker and each vowel in V1 position.

Speaker	V1	Formant	Point-biserial correlation
S1	e	ΔF_1	$r_{\mathrm{pb}} = 0.135$, NS, $N = 31$
	e	ΔF_2	$r_{\mathrm{pb}} = 0.132$, NS, $N = 31$
	o	ΔF_1	$r_{\mathrm{pb}} = 0.101$, NS, $N = 51$
	o	ΔF_2	$r_{\mathrm{pb}} = -0.016$, NS, $N = 51$
S2	e	ΔF_1	$r_{\mathrm{pb}} = -0.035$, NS, $N = 33$
	e	ΔF_2	$r_{\mathrm{pb}} = 0.143$, NS, $N = 33$
	o	ΔF_1	$r_{\mathrm{pb}} = 0.113$, NS, $N = 51$
	o	ΔF_2	$r_{\mathrm{pb}} = -0.310$, $p < 0.05$, $N = 51$
S3	e	ΔF_1	$r_{\mathrm{pb}} = -0.219$, NS, $N = 33$
	e	ΔF_2	$r_{\mathrm{pb}} = 0.206$, NS, $N = 33$
	o	ΔF_1	$r_{\mathrm{pb}} = 0.007$, NS, $N = 51$
	o	ΔF_2	$r_{\mathrm{pb}} = -0.276$, $p < 0.05$, $N = 51$
S4	e	ΔF_1	$r_{\mathrm{pb}} = 0.139$, NS, $N = 33$
	e	ΔF_2	$r_{\mathrm{pb}} = 0.201$, NS, $N = 33$
	o	ΔF_1	$r_{\mathrm{pb}} = -0.036$, NS, $N = 51$
	o	ΔF_2	$r_{\mathrm{pb}} = -0.072$, NS, $N = 51$
S5	e	ΔF_1	$r_{\mathrm{pb}} = -0.018$, NS, $N = 33$
	e	ΔF_2	$r_{\mathrm{pb}} = -0.176$, NS, $N = 33$
	o	ΔF_1	$r_{\mathrm{pb}} = 0.259$, NS, $N = 51$
	o	ΔF_2	$r_{\mathrm{pb}} = -0.181$, NS, $N = 51$
S6	e	ΔF_1	$r_{\mathrm{pb}} = -0.193$, NS, $N = 33$
	e	ΔF_2	$r_{\mathrm{pb}} = 0.199$, NS, $N = 33$
	o	ΔF_1	$r_{\mathrm{pb}} = 0.175$, NS, $N = 51$
	o	ΔF_2	$r_{\mathrm{pb}} = 0.114$, NS, $N = 51$

Table VII: Relationship between the average lexical frequency across the two members of each word pair and the size of vowel harmony effects for each speaker and each vowel in V1 position.

Speaker	V1	Formant	Kendall's tau
S1	e	ΔF_1	$\tau = 0.026$, NS, $N = 42$
		ΔF_2	$\tau = 0.149$, NS, $N = 42$
	o	ΔF_1	$\tau = 0.017$, NS, $N = 65$
		ΔF_2	$\tau = 0.025$, NS, $N = 65$
S2	e	ΔF_1	$\tau = 0.047$, NS, $N = 42$
		ΔF_2	$\tau = -0.155$, NS, $N = 42$
	o	ΔF_1	$\tau = -0.122$, NS, $N = 65$
		ΔF_2	$\tau = -0.079$, NS, $N = 65$
S3	e	ΔF_1	$\tau = 0.121$, NS, $N = 42$
		ΔF_2	$\tau = -0.074$, NS, $N = 40$
	o	ΔF_1	$\tau = 0.017$, NS, $N = 65$
		ΔF_2	$\tau = 0.175$, $p < 0.05$, $N = 65$
S4	e	ΔF_1	$\tau = -0.006$, NS, $N = 42$
		ΔF_2	$\tau = -0.104$, NS, $N = 42$
	o	ΔF_1	$\tau = 0.068$, NS, $N = 65$
		ΔF_2	$\tau = -0.090$, NS, $N = 65$
S5	e	ΔF_1	$\tau = -0.057$, NS, $N = 42$
		ΔF_2	$\tau = 0.062$, NS, $N = 42$
	o	ΔF_1	$\tau = -0.055$, NS, $N = 65$
		ΔF_2	$\tau = 0.014$, NS, $N = 65$
S6	e	ΔF_1	$\tau = 0.018$, NS, $N = 42$
		ΔF_2	$\tau = -0.094$, NS, $N = 42$
	o	ΔF_1	$\tau = -0.212$, $p < 0.01$, $N = 65$
		ΔF_2	$\tau = -0.178$, $p < 0.05$, $N = 65$

Table VIII: t and p values associated with each predictor in each linear mixed-effects model. See text for details.

Speaker	V1	Predictor	V2			
			e/ɛ	ø/œ	o/ɔ	i/a
S1	e	V2 height	$t(95) = 4.79$, $p < 0.001$	$t(51) = 1.39$, NS	$t(39) = 1.76$, NS	$t(81) = 0.49$, NS
		F_1 in V2	$t(95) = 4.18$, $p < 0.001$	$t(51) = 1.52$, NS	$t(39) = 2.59$, $p < 0.05$	$t(81) = -0.84$, NS
	o	V2 height	$t(179) = 3.03$, $p < 0.01$	$t(128) = 2.86$, $p < 0.01$	$t(25) = 0.95$, NS	$t(107) = -0.33$, NS
		F_1 in V2	$t(179) = 3.88$, $p < 0.01$	$t(128) = 3.58$, $p < 0.001$	$t(25) = 1.12$, NS	$t(107) = 0.58$, NS
S2	e	V2 height	$t(95) = -0.71$, NS	$t(53) = -0.51$, NS	$t(39) = -0.53$, NS	$t(95) = 0.63$, NS
		F_1 in V2	$t(95) = 2.08$, $p < 0.05$	$t(53) = 0.58$, NS	$t(39) = -0.45$, NS	$t(95) = 0.99$, NS
	o	V2 height	$t(179) = 0.86$, NS	$t(130) = 0.29$, NS	$t(25) = 0.25$, NS	$t(107) = -0.5$, NS
		F_1 in V2	$t(179) = 2.51$, $p < 0.05$	$t(130) = 1.99$, $p < 0.05$	$t(25) = 2.13$, $p < 0.05$	$t(107) = 0.75$, NS
S3	e	V2 height	$t(95) = 3.47$, $p < 0.001$	$t(53) = 1.54$, NS	$t(39) = 3.48$, $p < 0.01$	$t(95) = -0.11$, NS
		F_1 in V2	$t(95) = 2.38$, $p < 0.05$	$t(53) = 1.93$, NS	$t(39) = 1.69$, NS	$t(95) = -1.06$, NS
	o	V2 height	$t(179) = 1.49$, NS	$t(130) = 1.65$, NS	$t(25) = 1.58$, NS	$t(109) = 0.83$, NS
		F_1 in V2	$t(179) = 1.39$, NS	$t(130) = 1.83$, NS	$t(25) = 0.86$, NS	$t(109) = 2.4$, $p < 0.05$
S4	e	V2 height	$t(95) = 0.23$, NS	$t(53) = 3.42$, $p < 0.01$	$t(39) = 2.9$, $p < 0.01$	$t(95) = 0.42$, NS
		F_1 in V2	$t(95) = 0.07$, NS	$t(53) = 2.94$, $p < 0.01$	$t(39) = 2.29$, $p < 0.05$	$t(95) = 2.39$, $p < 0.05$
	o	V2 height	$t(179) = 1.79$, NS	$t(130) = 3.86$, $p < 0.001$	$t(25) = 1.89$, NS	$t(109) = 1.19$, NS
		F_1 in V2	$t(179) = 2.52$, $p < 0.05$	$t(130) = 3.55$, $p < 0.001$	$t(25) = 1.56$, NS	$t(109) = 0.4$, NS
S5	e	V2 height	$t(95) = 1.01$, NS	$t(53) = -0.11$, NS	$t(39) = 2.05$, $p < 0.05$	$t(95) = -0.64$, NS
		F_1 in V2	$t(95) = 2.87$, $p < 0.01$	$t(53) = 0.9$, NS	$t(39) = 1.25$, NS	$t(95) = -1.59$, NS
	o	V2 height	$t(179) = -1.03$, NS	$t(130) = -0.03$, NS	$t(25) = -1.46$, NS	$t(109) = -0.61$, NS
		F_1 in V2	$t(179) = 3.8$, $p < 0.001$	$t(130) = 1.38$, NS	$t(25) = 1.57$, NS	$t(109) = 3.05$, $p < 0.01$
S6	e	V2 height	$t(95) = -0.15$, NS	$t(53) = 0.33$, NS	$t(39) = 1.83$, NS	$t(95) = 0.6$, NS
		F_1 in V2	$t(95) = -1.86$, NS	$t(53) = 1.24$, NS	$t(39) = 0.75$, NS	$t(95) = 1.42$, NS
	o	V2 height	$t(177) = 1.5$, NS	$t(130) = 0.97$, NS	$t(25) = -0.87$, NS	$t(109) = 2.7$, $p < 0.01$
		F_1 in V2	$t(177) = 1.25$, NS	$t(130) = -1.21$, NS	$t(25) = -0.94$, NS	$t(109) = 2.68$, $p < 0.01$

Figure legends

Figure 1: Average values of F_2 and F_1 for each vowel and each speaker. 1-sigma ellipses encompassing ±1 standard deviation along the two principal components for each vowel are also shown.

Figure 2: Average F_2 and F_1 frequencies for front and back V1s /e/ and /o/ in Northern Speaker S1. Each panel corresponds to one of the four V2 pairs. Open circles: before a low V2; filled circles: before a non-low V2. The open circle and filled circle associated with each of the two V1s are connected to each other with a line segment. The average F_2 and F_1 frequencies for /i, e, ɛ, a, ɔ, o/ in V2 position are also shown as a reference grid in grey.

Figure 3: Average F_2 and F_1 frequencies for front and back V1s /e/ and /o/ in Northern Speaker S2. Open circles: before a low V2; filled circles: before a non-low V2.

Figure 4: Average F_2 and F_1 frequencies for front and back V1s /e/ and /o/ in Northern Speaker S3. Open circles: before a low V2; filled circles: before a non-low V2.

Figure 5: Average F_2 and F_1 frequencies for front and back V1s /e/ and /o/ in Southern Speaker S4. Open circles: before a low V2; filled circles: before a non-low V2.

Figure 6: Average F_2 and F_1 frequencies for front and back V1s /e/ and /o/ in Southern Speaker S5. Open circles: before a low V2; filled circles: before a non-low V2.

Figure 7: Average F_2 and F_1 frequencies for front and back V1s /e/ and /o/ in Southern Speaker S6. Open circles: before a low V2; filled circles: before a non-low V2.

Figure 8: Relationship between F_1 frequency in V2 vs V1 for low V2s (open circles) and non-low V2s (filled circles); (a) Speaker S1, back V1 before ø/œ; (b) Speaker S3, front V1 before o/ɔ; (c) Speaker S5, back V1 before e/ɛ. The regression line associated with each of the two V2 categories is also shown.

49

Figure 1:

Speaker S1
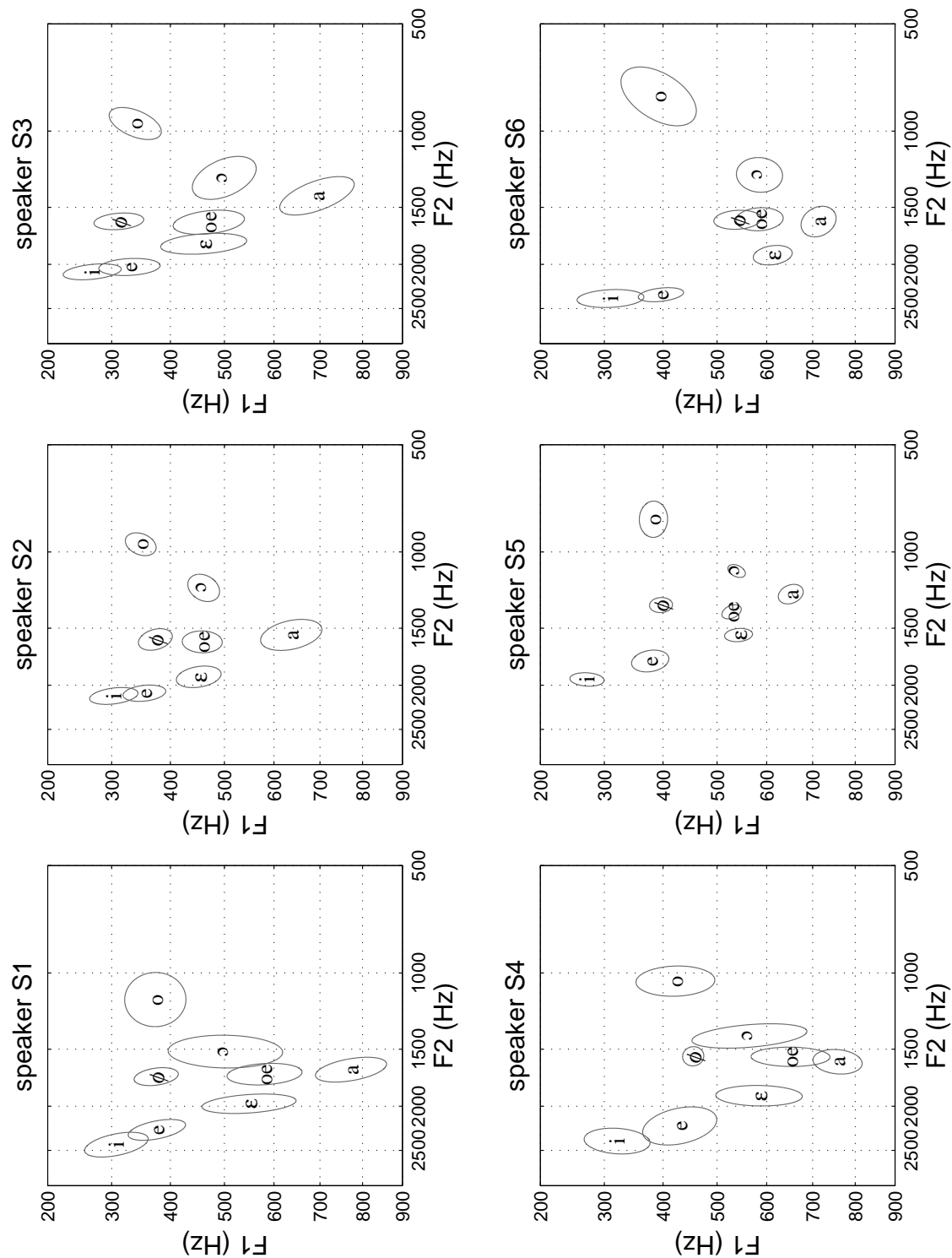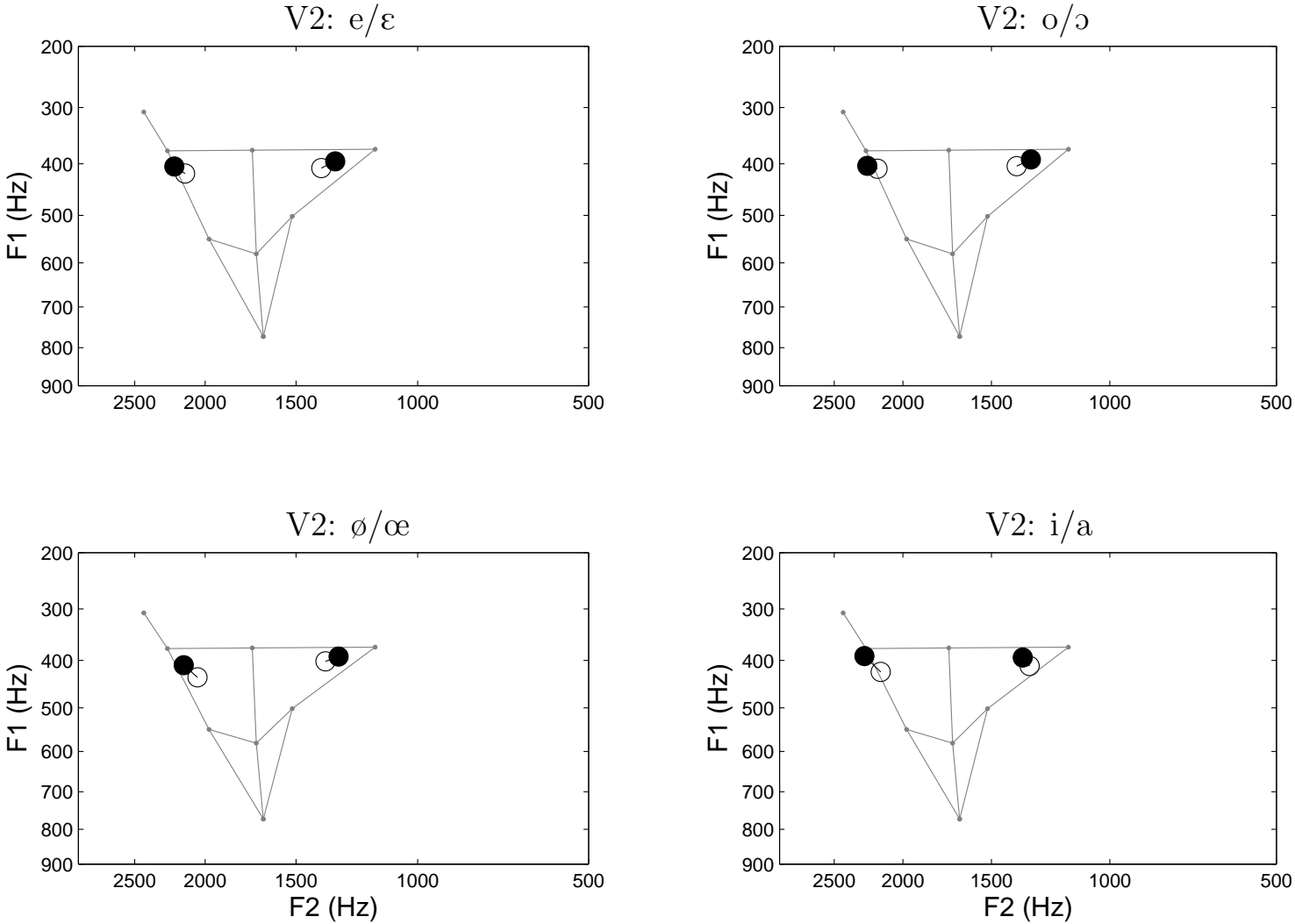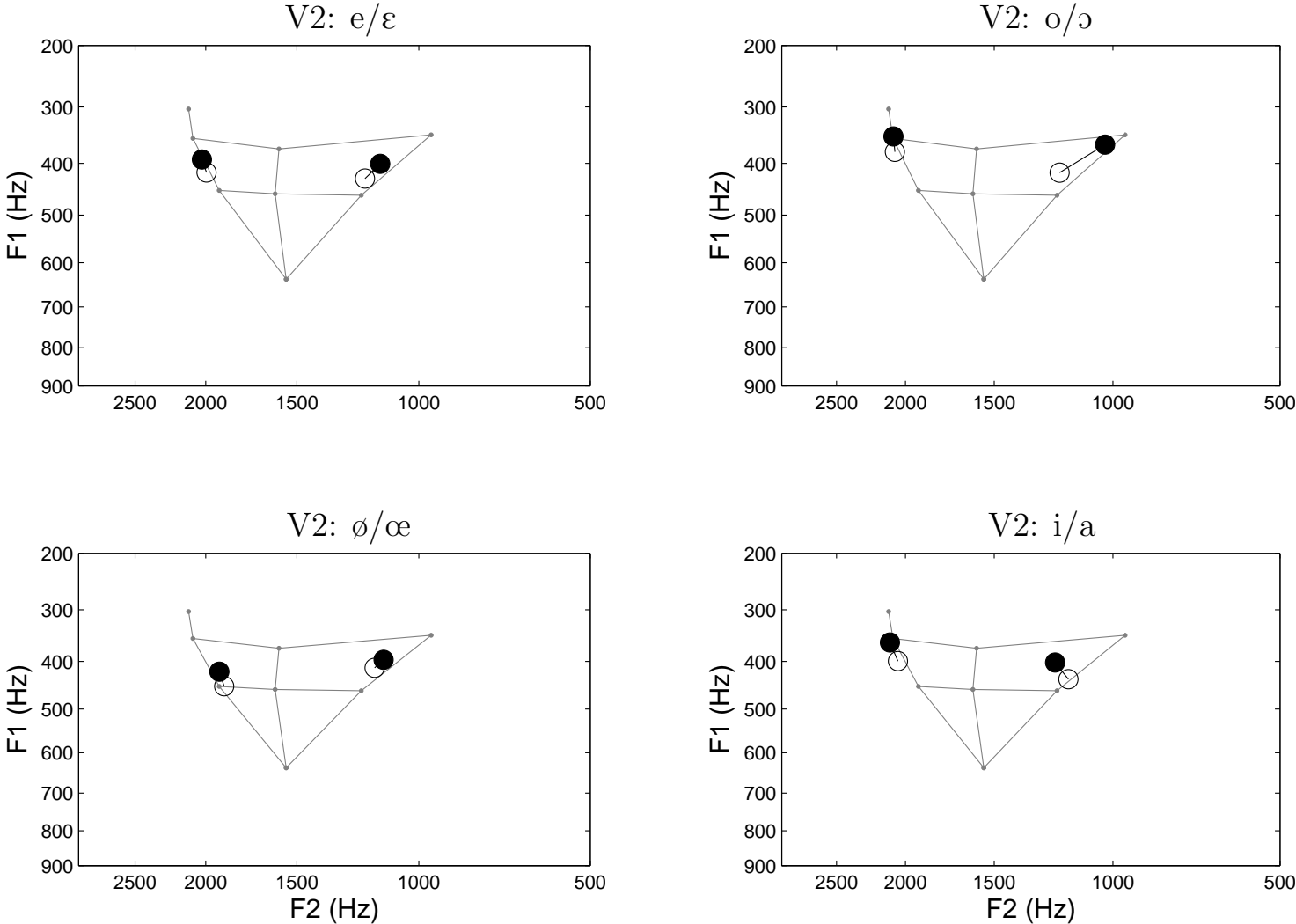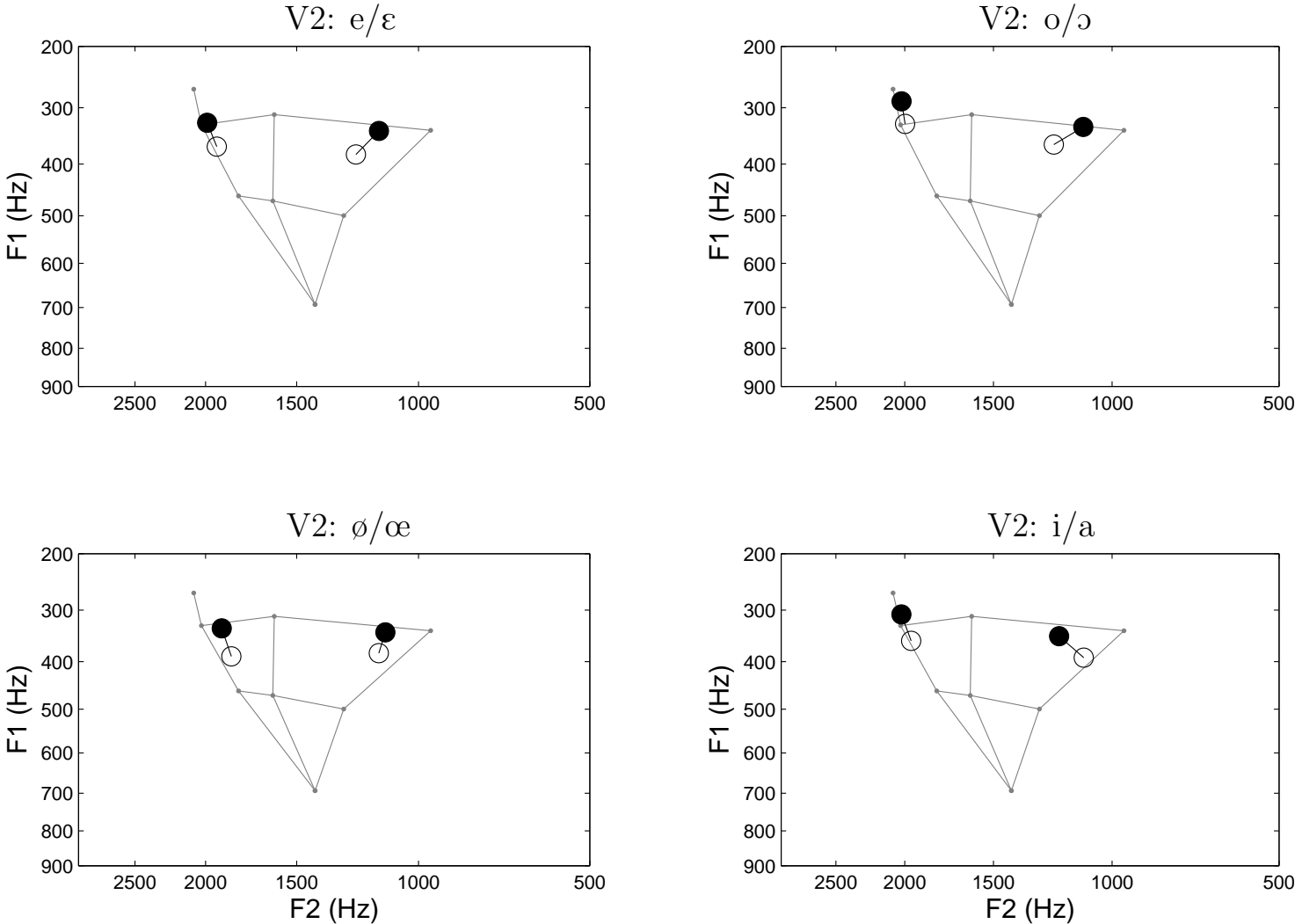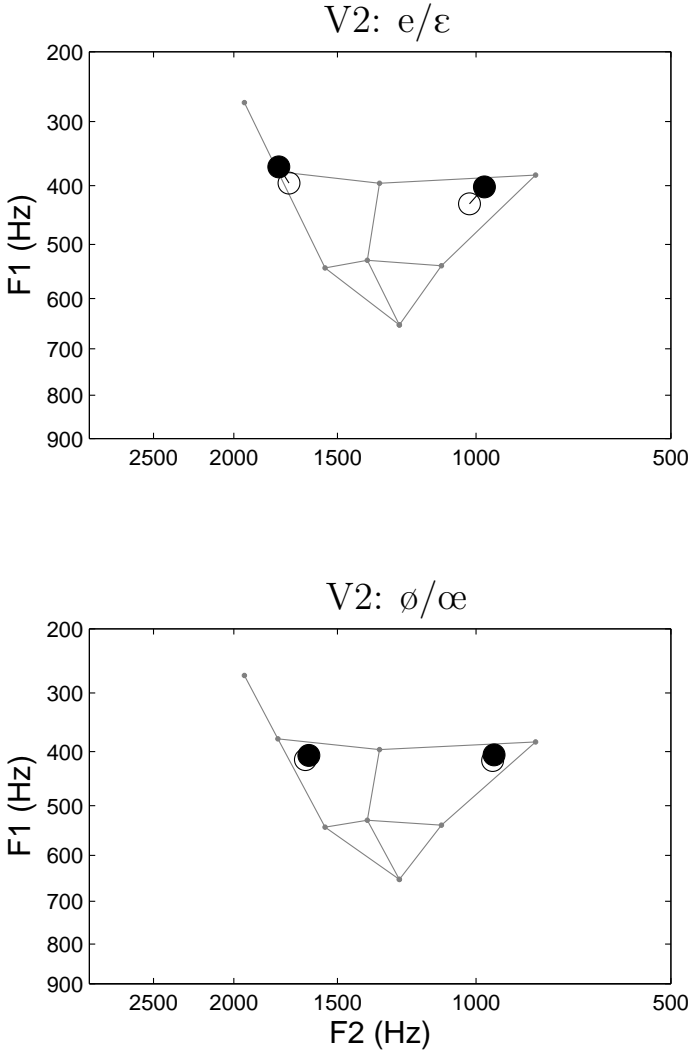

Figure 2:

Speaker S2

V2: e/ɛ

V2: o/ɔ

V2: ø/œ

V2: i/a

F1 (Hz)

F2 (Hz)

Speaker S3

Speaker S4

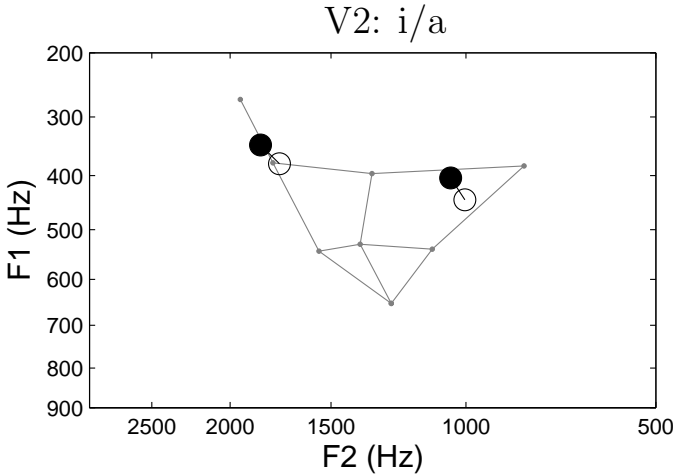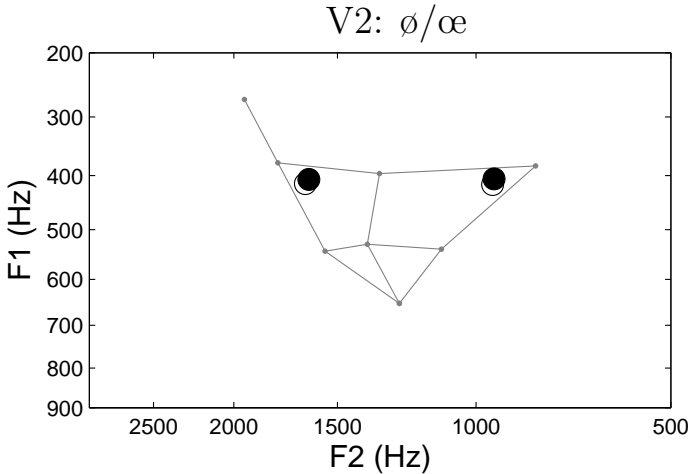V2: e/ɛ   V2: o/ɔ   V2: ø/œ   V2: i/a

F1 (Hz)

F2 (Hz)

54
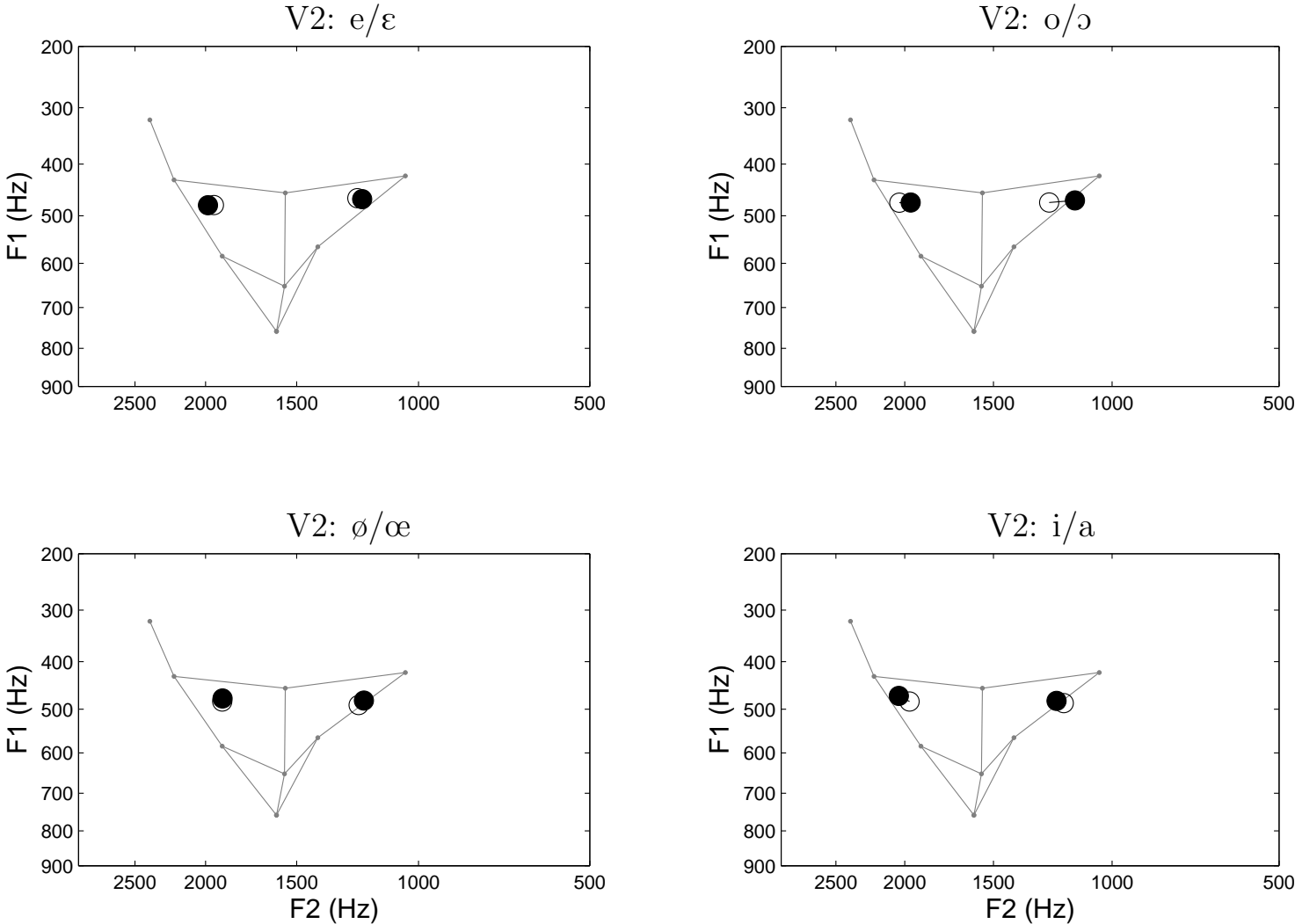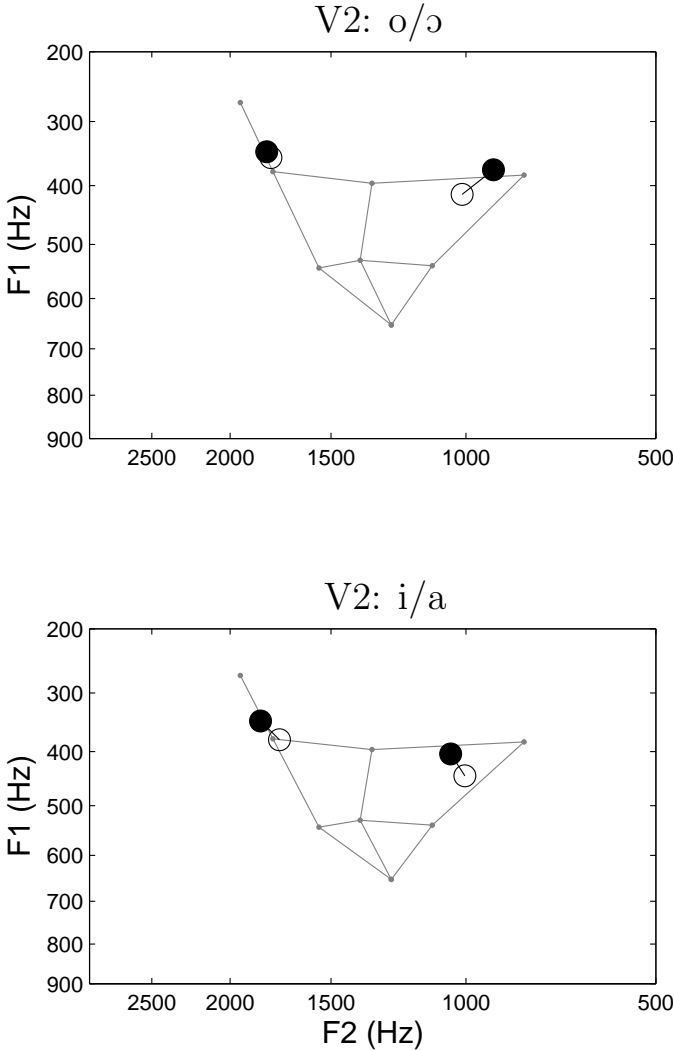
Figure 5:

Speaker S5

Speaker S6

Figure 7:

Figure 8: