

## Variation in the lexical distribution and implementation of phonetically similar phonemes in Catalan

Marianna Nadeu<sup>a\*</sup>, Margaret E. L. Renwick<sup>b</sup>

\*Corresponding author: [nadeu@psu.edu](mailto:nadeu@psu.edu), Tel: +1 814 865 4252

<sup>a</sup> Penn State, Burrowes Building, University Park PA 16803, USA / [nadeu@psu.edu](mailto:nadeu@psu.edu)

<sup>b</sup> University of Georgia, 240 Gilbert Hall, Athens GA 30602, USA / [mrenwick@uga.edu](mailto:mrenwick@uga.edu)

### Abstract

In some Romance languages with two pairs of mid vowel phonemes, it is acknowledged that these contrasts are somewhat unstable. We analyze the distribution and realization of the anterior and posterior mid vowels in Catalan to test claims (mostly based on anecdotal evidence) that these contrasts exhibit inter- and intraspeaker variability. Participants produced target words containing stressed mid vowels and, later, judged vowel height (/e/ vs. /ɛ/; /o/ vs. /ɔ/) in the same words. The results indicate that, even intradialectally, the distribution of mid vowels is somewhat variable, with speakers showing only moderate agreement in the distribution of phonemic vowels. In addition, speakers are not always consistent in their realization of mid vowels when they produce the same word (probably indicating weak phonolexical representations). Interspeaker variation was also observed in the phonetic implementation of the contrasts. The results indicate that the Catalan mid vowel contrasts, like those in other Romance languages, are weaker and less stable than other phonological oppositions.

### Highlights

- The lexical distribution of Catalan mid vowels exhibits individual variation.
- The phonetic implementation of these vowels also shows individual variation.
- Speaker intuitions of mid vowel quality do not always match their production.
- Adjacent mid vowels have more acoustic overlap than other adjacent vowel pairs.
- Some phonemic contrasts are less robust than others.

### Keywords

Mid vowels; phonological contrast; Romance languages; individual variation; Catalan; acoustic overlap; bilingualism

## Variation in the lexical distribution and implementation of phonetically similar phonemes in Catalan

### 1. Introduction

In this paper, we investigate the strength of the phonological contrasts between two pairs of vowel phonemes, as realized by individual speakers of Central Catalan. Like other Romance languages including French and Italian, many varieties of Catalan have two pairs of mid vowels: anterior /e, ε/ and posterior /o, ɔ/ (Recasens, 2014, p. 15). As indicated by our symbolic representation, these four vowels are all assumed to be separate phonemes, meaning that they lie in “direct contrast” with one another, distinguishing the meanings of words (Bloch & Trager, 1942, p. 38; Bloomfield, 1933, pp. 77–79). That is, there exist minimal pairs distinguished only by [e] vs. [ε], or [o] vs. [ɔ] (for example, /be/ ‘good’ vs. /bε/ ‘lamb’; /os/ ‘bear’ vs. /ɔs/ ‘bone’).

However, the phonological status of mid vowels in Catalan is complex in several ways. First, mid vowel usage varies across dialects. For instance, a word with [e] in one variety may be realized with [ε] in another. Second, impressionistic descriptions point to oscillation in mid vowel assignment even within the same dialect or by the same speaker. Third, in Catalonia, no monolingual speakers of Catalan exist. All speakers, even those most dominant in Catalan, are bilingual in Spanish, whose vowel system lacks /ε, ɔ/. Thus, the realization of the contrasts may depend on a speaker’s exposure to Spanish (Mora, Keidel, & Flege, 2015; Mora & Nadeu, 2012). The combined effects of these factors motivate the focus of this paper, on individual speaker differences in the realization of the /e – ε/ and /o – ɔ/ contrasts. We investigate these differences using a method that pairs the phonetic analysis of speakers’ productions with their native intuitions of vowel quality, accessed at the word level.

#### 1.1 Degrees of phonological contrast

A central question in phonological analysis is whether a given pair of sounds lie in opposition (e.g. Trubetzkoy, 1939), and the traditional test is to seek minimal pairs: if at least one is found, then the two sounds in question are separate phonemes. If, on the other hand, the two sounds appear in predictably complementary distribution, they are allophones of the same phoneme, and are not phonologically contrastive (Hockett, 1955; Hyman, 1975, sec. 3.1). While the notion of phonemic categories remains highly useful in theoretical linguistics and related fields (Hall, 2013; Ladd, 2006), the idealized view of phonemic status based solely on lexical contrast is increasingly called into question. In their discussion of quasi-phonemic contrasts in Scottish English, Scobbie & Stuart-Smith (2008) make clear that not all contrasts are equally contrastive. Indeed, some contrasts are supported by many minimal pairs and salient phonetic distinctions, while others have a low functional load (Hockett, 1966; Renwick, 2011; Surendran & Niyogi, 2006), making them more susceptible to merger over time (Wedel, Jackson, & Kaplan, 2013). Some contrasts are phonetically overlapping, such as palatal glides vs. obstruents in Castilian Spanish (Hualde, 2004) or the alveolar/retroflex contrast in Arrernte (Tabain, 2009). Other weak contrasts may be partially phonologically conditioned. For instance, in English, /ð/ and /θ/ are historical allophones that gained contrastiveness due to the lenition of word-initial [θ] to [ð] in function words (*the, thy*), coupled with the appearance of [ð] under stress (Dobson, 1968; Luick, 1940; Thurber, 2011). Nonetheless, few minimal pairs exist (e.g. *either – ether*). As reviewed by Hall (2013), there are many causes of these *intermediate phonological relationships* or *marginal contrasts*. They are common across languages and have been given numerous labels.

Marginal contrasts occur among vowels across the Romance languages. In Romanian, the phonemic contrast between /i, ʌ/ has a low functional load maintained by only a handful of minimal pairs; otherwise, these historical allophones remain in nearly complementary distribution (Renwick, 2014). Turning to Romance languages with /e, ε/ and /o, ɔ/, several authors have noted the instability of these contrasts, variability in their realizations, neutralization in certain varieties, or a different behavior with respect to other vowel contrasts (for French: Fagyal, Kibbee, & Jenkins, 2006; Landick, 1995; Trubetzkoy, 1939; for Italian: Renwick & Ladd, in preparation; for Portuguese: Escudero, Boersma, Rauber, & Bion, 2009; Wetzels, 1992). The fact that not all contrasts are equal is buoyed by recent work on Romanian (Renwick, 2014) and French (Hall & Hume, 2015), showing that marginal contrasts among vowels are less perceptually distinct than more robust contrasts, even when the segments are acoustically distinct.

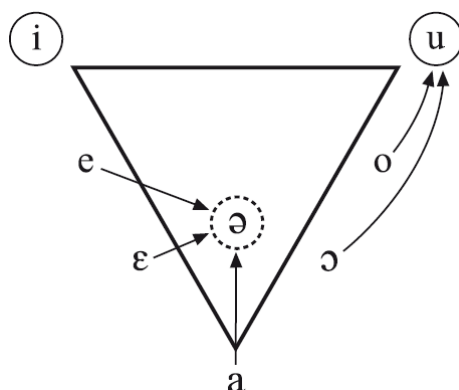
The mid vowel distinctions in Catalan, similarly, are demonstrably weaker than other vowel contrasts in the language. Some researchers have anecdotally observed variation involving these pairs of vowels and have compiled lists of words that seem to be prone to vowel substitutions. In addition, based on the results of psycholinguistic experiments, Mora, Keidel, and Flege (2015) hypothesized that, due to language internal and external factors, the mid vowel contrasts are phonologically weakened, even in Catalan-dominant bilinguals.

Regardless of the source of this variation, the primary focus of the present study is on whether speakers' intuitions match their production, a correspondence not previously investigated for Catalan. Our study, with respect to earlier work, begins with a methodological acknowledgment that individual speakers vary in assigning vowel height to particular lexical items. While previous studies have demonstrated effects of language dominance on production and perception in Catalan, they have not taken into account the possibility that two speakers, even with equivalent Catalan dominance levels, may possess different lexical representations for the same word. In fact, the mapping between individual words and the category in which they are realized is rarely considered. An exception is Bosch and Ramon-Casas' (2011) study, which in calculating a production "error rate" assumes a "correct" phonological form to which speakers should adhere to achieve native-like speech. Instead, we probe each speaker's phonological intuitions, to create an individual standard of comparison based on the speaker's judgments. Against this backdrop, we evaluate the strength of Catalan mid vowel contrasts in speech production. We hypothesize that if speakers' mid vowel contrasts are robust, then their productions will match their native intuitions of vowel quality; and that vowel tokens from repetitions of a single lexical item will cluster together acoustically. If mismatches are found, either between speakers' production and intuition or across repetitions of a word, we take this as evidence of fuzziness, or weakness, in the phonological category boundaries between Catalan mid vowels.

In the remainder of the Introduction, we provide a brief historical sketch of Catalan mid vowels, their synchronic distribution, and review previous experimental work on which our study builds. The rest of the article is organized as follows: section 2 summarizes the goals and hypotheses. Section 3 describes the methodology. Section 4 presents the results, while the discussion and conclusions are found in section 5.

### 1.2 Origin and distribution of the mid vowels

Like Vulgar Latin (VL; Alkire & Rosen, 2010, p. 13), Central Catalan displays a seven-vowel system with two pairs of mid vowels (Figure 1)<sup>1</sup>, although these seven vowels are only possible in stressed syllables. Due to the existence of stress-conditioned vowel alternations, only three vowels, [i ə u], are possible in unstressed syllables (Mascaró, 2002). Despite the similarity between Latin and Catalan in their stressed vowel systems, the evolution of the anterior mid vowels is actually quite complex.



**Figure 1** *The vowel system of Central Catalan. The arrows show patterns of phonological vowel reduction. Circled vowels are those that can appear in stressed and unstressed syllables. Schwa can only appear in unstressed syllables.*

Fabra (1906) observed that VL /e/ (< Ē, Ī, OE) lowered to present-day Catalan /ɛ/ (with an intermediate step /ə/, preserved only in some varieties of Majorcan Catalan; Rasico 1993), whereas /ɛ/ (< Ē, AE) evolved into /e/, with some exceptions.<sup>2</sup> Specifically, /ɛ/ was preserved unchanged before /l/, /r/ + vowel, /r/ + non-labial consonant, [w], and before the VL group N'R (see also Rasico 1981).<sup>3</sup> This crossover evolution exhibited by the anterior mid vowels is not shared by the posterior series (Badia Margarit 1981: 145ff, Moll 1991:80-81, Gulsoy 1993; although see Coromines 1974: 189): VL /ɔ/ > Central Catalan /ɔ/, and VL /o/ > Central Catalan /o/.

Despite the mid vowels forming phonemic contrasts, some patterns or tendencies in the distribution of mid vowels have been identified. That is, phonological context affects the distribution of mid vowels in Catalan. First (and partially as predicted from the historical

<sup>1</sup> To simplify, we focus on the Central variety of Eastern Catalan only. While the seven-vowel system is quite common across Catalan varieties, smaller and larger vowel systems are found (e.g. Majorcan Catalan has eight stressed vowels, Girona Catalan has six, and Roussillonais Catalan has only five; Recasens, 2014, p. 15). In addition, the diachronic evolution and, therefore, contemporary synchronic distribution of mid vowels is not shared by all dialects of Catalan (Badia i Margarit, 1981, p. 138; Moll, 1991, p. 72ff; Wheeler, 2005, p. 38). Catalan varieties also differ importantly in their patterns of phonological vowel reduction (Mascaró, 2002).

<sup>2</sup> The change from Vulgar Latin /e/ to Central Catalan /ɛ/ is not without exceptions, but these seem to lack a systematic character (Rasico, 1981).

<sup>3</sup> N'R refers to sequences of /n/ and /r/ that originated due to the syncope of an intervening vowel. These new sequences tended to be repaired as /ndr/ (Mascaró, 2008), as in GÈNĒRU > *gendre* 'son-in-law'.

evolution of the front mid vowels discussed above), /e/ is said to not appear before [l], [r] followed by a coronal consonant or a vowel, [ŋ], [ɲ], [f], and [sk] (Mascaró, 2008; Wheeler, 2005, p. 39). Going beyond the immediate context, Mascaró (1984, 2008, 2010) and Wheeler (2005) attempted to establish other factors that condition the distribution of mid vowels. Low mid vowels are triggered by the presence of prestressed suffixes (Mascaró, 1984), posttonic /i, j/ and (much less strongly) /u/ (Mascaró, 2008; Wheeler, 2005). In addition, Catalan speakers are claimed to favor low mid vowels in their pronunciation of borrowings, neologisms, cultisms, or brand names (Badia Margarit, 1966; Mascaró, 2008; Pi-Mallarach, 2001, 2006, 2007), although a strong tendency to produce loanwords in ways that depart from the native phonology has also been attested (i.e., without low mid vowels and without vowel reduction, Cabré, 2009).

### *1.3 Present-day variation in the realization of mid vowels*

Attempts to formally account for the distribution of mid vowels through phonological rules (Mascaró, 1984, 2008, 2010; Wheeler, 2005) in a categorical manner, or to quantify their frequency in the lexicon (Rafel i Fontanals, 1980), stand in stark contrast with claims, as early as the beginning of the 20<sup>th</sup> century, that mid vowels are realized variably within the same dialect and geographical area.<sup>4</sup> Although, as we noted, there are patterns in the etymological correspondences between Latin and Catalan vowels, synchronic reality for native speakers is much more variable.

In addition to phonological conditioning, Fabra (1906) noted variable realization of the anterior mid vowels in cultisms and, less frequently, in “words of popular origin” (p. 23). Badia i Margarit (1968, 1969a, 1969b, 1970a, 1970b) reported variation in the realization of the anterior and posterior mid vowels, such that the very same word might be produced with the low or high mid vowel by different speakers of the same region (usually Barcelona) or even by the same speaker.<sup>5</sup> He provided examples from oral speech and literary texts. In particular, he closely analyzed rhymes of 19<sup>th</sup> and 20<sup>th</sup> century poems and noted that most authors allowed low mid vowels to rhyme with high mid vowels (Badia i Margarit, 1968, 1970a; see also Coromines [1974, p. 219], who noted the same phenomenon in medieval poetry). Moll (1991) claimed that *Renaixença* (19<sup>th</sup> century) poets rhymed /o/ and /ɔ/ words because their rhymes were “visual, rather than acoustic” (p. 80), owing to their training in Spanish literary norms.<sup>6</sup> Yet, imperfect rhymes mixing high and low mid vowels are not only found in written texts, but also in traditional (Badia i Margarit, 1969b) and contemporary songs, where visual rhymes are more difficult to justify.

---

<sup>4</sup> Wheeler (2005, pp. 45–46) acknowledged the existence of individual variation, which, he noted, is a source of instability of the mid vowel contrasts, together with the “major dialect difference”.

<sup>5</sup> Some of the articles by Badia i Margarit show different years of publication although they appear in the same edited volume. The author edited a volume that reprinted articles and book chapters that he had published in other venues previously. We refer to the original years of publication to allow the reader to see Badia i Margarit’s progression of thought and facilitate keeping track of his different studies, but the references include the edited volume where all the works can be more easily found.

<sup>6</sup> There is one single orthographic symbol for the front and another one for the back mid vowels (*e* represents both /e/ and /ɛ/, while *o* represents the two back vowels). When an orthographic stress mark is required, the height of the mid vowels is then represented orthographically (*è* = /ɛ/, *é* = /e/, *ò* = /ɔ/, *ó* = /o/).

Recasens (1993) acknowledged that certain words must be afforded two normative pronunciations. In trying to account for the range of words and suffixes admitting variation between low and high mid vowels, he mentioned analogy; influence from Spanish (but see Badia i Margarit, 1969b, 1970b); and the existence of phonological vowel reduction. The stress-final infinitives (e.g. *arrencar* [əɾəŋ'ka] 'to pluck', *galopar* [gəlu'pa] 'to gallop') do not provide information on the identity of the vowel in root-stressed forms, which may generate doubts for the speaker (is the third person singular present indicative form [ə'reŋkə] or [ə'reŋkə]; [gə'ləpə] or [gə'ləpə]?<sup>7</sup>), as Badia i Margarit (1970b) had already discussed. Badia i Margarit (1970b) also claimed that confusability of the mid vowels may be increased by the fact that certain nouns may not present the same vowel as the morphologically related verb (*pes* ['pes] 'weight' vs. *peso* ['pezu] '(I) weigh'). Yet another important source of confusion or variation is the existence of extensive interdialectal differences (Mora & Nadeu, 2012; Wheeler, 2005), due, for example, to the fact that Vulgar Latin /e/ was preserved as such in North-Western Catalan, Valencian, and Alguerese (Rasico, 1981). However, Badia i Margarit (1970b) cast doubt on this hypothesis. In summary, the outcome of VL mid vowels is subject to certain dialect-specific phonological constraints in Catalan; but additionally the literature describes sources for and anecdotes of variation in their realization, regardless of dialect or phonological context.

#### *1.4 Perception and production of the mid vowel contrasts*

A plethora of psycholinguistic work has capitalized on the Catalan mid vowels to examine plasticity in the perception and production of non-native contrasts. A full review of this work is beyond the scope of this paper: we discuss relevant findings from studies of adults speaking Central Catalan, focusing on the performance of Catalan-dominant bilinguals.<sup>8</sup> Most of this work has focused on the perception of the /e – ε/ contrast.

Perception studies show that L1 Catalan bilinguals perceive the mid vowels reliably, unlike L1 Spanish Catalan-Spanish bilinguals, whose performance with these contrasts has been interpreted as a seriously compromised ability to acquire new phonological categories in the L2, despite early and intensive exposure. Catalan-dominant participants' performance showed an S-shaped curve in an identification task and a peak in a discrimination task (Pallier, Bosch, & Sebastián-Gallés, 1997), and decreased discrimination around prototypical realizations (Bosch, Costa, & Sebastián-Gallés, 2000). In addition, hearing one member of a minimal pair did not prime recognition of the other member (Pallier, Colomé, & Sebastián-Gallés, 2001); and alternation between /e/ and /ε/ in the second syllable of nonce words slowed down participants' classification of the first syllable of those words (Navarra et al. 2005), indicating accurate perception of the contrast.

Nevertheless, some details in these studies hint at the possibility that Catalan dominant speakers' perception of the mid vowels may be less robust than for other contrasts, that asymmetries in the behavior of mid vowels exist, and that there is lexical variation. In Pallier et al. (1997), the S-

<sup>7</sup> These are some of Recasens's (1993) examples of words that admit both mid vowels.

<sup>8</sup> Our review does not include studies on infant acquisition of the mid vowel contrasts (see e.g., Bosch & Sebastián-Gallés, 2003; Sebastián-Gallés & Soto-Faraco, 1999) or those focusing on Majorcan Catalan (e.g., Amengual, 2014; Simonet, 2011, 2014), owing to the important sociolinguistic differences between the Central and Majorcan Catalan contexts and significant differences in the historical evolution and acoustic realization of the mid vowels.

shaped curve was compressed: the extremes of the continuum did not reach 0 and 100% of /ε/, but rather approached 15% and 80% respectively. Mora, Keidel and Flege (2011) confirmed that the /e – ε/ and /o – ɔ/ contrasts were perceived less categorically than /i – e/ and /u – o/.

As for the asymmetries, in Navarra et al. (2005), participants performed faster in the /e/ condition than in the /ε/ condition. In a lexical decision task involving /e – ε/ substitutions (Sebastián-Gallés, Echeverría and Bosch, 2005), e-type non-words (created by replacing an original /e/ with /ε/) received a correct response at approximately the same rate as real words, but the same was not true for ε-type non-words (below 60% correct responses). In Sebastián-Gallés, Rodríguez-Fornells, de Diego-Balaguer and Díaz (2006), N400 patterns to words did not differ from those elicited by non-words (failing to show the expected enhancement for non-words). In another study using the same tasks and stimuli (Sebastián-Gallés, Vera-Constán, Larsson, Costa, & Deco, 2009), e-type non-words did present larger N400 amplitude than e-type words (as expected), but no differences were found between ε-type words and non-words. The authors concluded that exposure to Spanish-accented Catalan (where /ε/ is frequently realized as /e/) results in Catalan-dominant speakers acquiring new lexical representations without altering their phonemic representations. In particular, ε-type words only would involve an “empty slot” (Sebastián-Gallés et al., 2009, p. 2353), some sort of mid vowel underspecified for height.

Evidence for the fact that not all words with mid vowels behave the same way comes from the lexical decision task in Sebastián-Gallés et al. (2009), where two e-type and one ε-type words (and their non-words) had to be discarded due to “high error rates” (p. 2346). The remaining words elicited very few errors (< 6%). A study with Catalan monolingual toddlers (Ramon-Casas, Swingley, Sebastián-Gallés, & Bosch, 2009) involving again words with /e – ε/ changes showed that 18 out of the 24 participants displayed greater fixation times to the correctly pronounced tokens than to mispronunciations, but not the other six participants. In addition, this was only true for three of the four target words included.

The production of mid vowel contrasts has not been studied so extensively. Bosch and Ramon-Casas (2011) showed that Catalan- and Spanish-dominant speakers phonetically implemented the /e – ε/ contrast quite similarly. Despite this, variation among Catalan-dominant speakers was observed in the distance between the vowels and in whether the contrast is mainly realized along F1 or also F2. In addition, three Catalan-dominant speakers had zero vowel selection errors (words not realized with the mid vowel that the researchers expected were computed as errors), while the remaining five had one or two (out of 26 words), showing again some lexical variation. The stimuli in that study were uncontrolled for phonological context, and errors were distributed across all words regardless of their characteristics. Mora and Nadeu (2012) found effects of Spanish L2 use on Catalan vowel perception and production by Catalan dominant speakers, which were greatest when speakers produced Catalan words with Spanish cognates.

In summary, even among Catalan-dominant speakers, some individual variation in production is observed, and perception of these contrasts is less robust than for other contrasts. Mora, Keidel and Flege (2015) proposed a phonological or perceptual weakening hypothesis, according to which both language internal and external factors (i.e. contact with and influence from Spanish) may be eroding the mid vowel contrasts in Catalan-dominant bilinguals. While we do not want to downplay the influence that Spanish may have or have had in the weakening of this contrast, the

presence of similar patterns in other Romance languages that are not in contact with Spanish support the idea that internal factors lie at the origin of these contrasts' instability. The current study sheds light on this phenomenon by exploring the production of both front and back mid vowel contrasts by Catalan-dominant Catalan-Spanish bilinguals. Although we do not address perception, we believe there may be a link between previous findings of weakened category boundaries and the presence of representational variability across speakers.

## 2. Goals and Hypotheses

This study investigates the acoustic realization of Central Catalan mid vowels in conjunction with native speakers' intuitions of vowel height. The goals of the paper are (1) to quantify the extent to which there is individual variation in the distribution of mid vowels in the lexicon, (2) to explore individual differences in the phonetic implementation of the mid vowel contrasts, and (3) to examine whether speakers' intuitions about mid vowel quality consistently match their acoustic realizations. These questions aim at investigating the strength of mid vowel contrasts in Central Catalan, and allow us to test claims that these contrasts are phonologically weak and that speakers' phonolexical representations involving mid vowels may be less stable than those comprising other vowels.

Because previous work has shown variation in the assignment of Catalan mid vowel quality, we believe that a study of individual differences must begin with a survey of speakers' intuitions. Along the way, we test the degree of interspeaker agreement that obtains when speakers consciously assign vowel quality to a series of words containing mid vowels. Because the mid vowels are phonologically contrastive and the speakers come from the same geographical area, we expect speakers to agree substantially in their assignment of mid vowel quality. However, we expect some amount of disagreement, based on claims that, within the same dialect, the very same word can be produced variably with a high or low mid vowel (see section 1.3).

Turning to the phonetic realization of the contrasts, we expect speakers to produce vowels that reflect their intuitions. For example, realizations of vowels that speakers label as /e/ should have the acoustic characteristics of a high mid front vowel. Leaving aside cases of vowel quality disagreement across speakers, it is predicted that speakers will implement the high vs. low mid vowel contrasts similarly.

It has been argued that one goal of phonological contrast selection and speech production is the enhancement of contrast distinctiveness (Flemming, 1995; Stevens & Keyser, 2010; but cf. Lindblom, 1986). Under this premise, one might expect vowels to be distributed in the vowel space in such a way that contiguous vowels are approximately equidistant in F1 (see Diehl, 2008). We thus investigate the amount of acoustic overlap between mid vowels and other pairs of neighboring vowels, hypothesizing that if speakers maintain distinct high and low mid vowels, then the amount of overlap between them should be no greater than that among other vowel pairs.

Our final goal probes the stability of phonolexical representations by exploring how consistent speakers are in producing a given vowel across repetitions of the same word. If speakers have clear intuitions of which vowel occurs in a certain word, that word should be uniformly and



exclusively pronounced with that vowel. Variable realization of the mid vowel in the same word may suggest unstable phonolexical representations.

### **3. Methods**

#### *3.1 Participants*

Given that all Central Catalan speakers are bilingual in Spanish, participants were asked to complete the Bilingual Language Profile Questionnaire (BLP; Birdsong, Gertken, & Amengual, 2012) online, as a prescreening test. The BLP questionnaire computes a language dominance score that ranges from  $-218$  to  $+218$ , based on participants' responses to a series of questions about the speaker's language history, use, proficiency, and attitudes. Values close to 0 represent balanced bilingualism. In this case, positive values indicate dominance in Catalan (the extreme, 218, would represent monolingualism in Catalan). 22 people (all female) completed the BLP questionnaire. We invited those participants who scored 60 or higher in the BLP (indicating Catalan dominance) to complete the experiment, resulting in 14 participants. These participants' mean score was 91.36 (range: 68.39–117.70). For comparison, the Catalan-dominant Majorcan participants in Amengual (2014) scored between 10 and 130 approximately, and those in Simonet (2014) had a mean BLP score of 22. The BLP does not include questions about participants' parents' place of birth or language. However, all our speakers confirmed that they had been exposed to Catalan from birth and that that language was spoken at home. Only four speakers (f02, f03, f05, and f13) reported using some Spanish (10% vs. 90% of Catalan) to communicate with their family. Note also that speakers did not know, ahead of time, that we would only be recruiting Catalan-dominant speakers. They were given the opportunity to complete the language background questionnaire in either Catalan or Spanish (all of our participants completed it in Catalan). The BLP questionnaire also does not include questions to determine participants' socioeconomic status. We thus cannot exclude the possibility that our speakers differ along these lines. However, we are not aware of any study indicating variation in Catalan mid vowel production that is conditioned by socioeconomic status.

All the participants were undergraduate or graduate students at the Universitat Rovira i Virgili (Tarragona, Spain). Some of them knew each other because they were members of the same social organization, and three were sisters (of these, two were twins). They were all born, raised, and currently living in the Central-Catalan-speaking area of the Camp de Tarragona region. Mean age was 20.83 (range: 18–27). Thus, the speakers formed a quite homogeneous group: they were all female, currently enrolled as students at the same university, mostly in their early 20s, and from the same region.

#### *3.2 Materials*

A set of 210 target items was created by manipulating stressed vowel, phonological context, word length, and (stressed) syllable structure. Target vowels (all seven vowels that form the Central Catalan stressed vowel system) were elicited in words with penultimate stress, from the first syllable of disyllabic and the second syllable of trisyllabic words. In the disyllabic words, the stressed syllable was either open or closed. To control for distant V-to-V coarticulation, the posttonic vowel was [ə] in all of the target words. In trisyllabic words, the pretonic vowel was also schwa, except in seven cases (due to lexical gaps). To minimize vowel quality variation due to phonetic context, the place of articulation of the preceding and following consonants was kept constant. The following consonant was always dentoalveolar. The preceding consonant was

labial in half of the target words and dentoalveolar in the other half, which allowed us to include a larger number of target words. Trills were included, although it is known that they tend to increase the preceding vowel's F1, lowering the vowel (Recasens, 1985). *Nasals and laterals were avoided as much as possible in order to avoid vowel nasalization (in the case of nasals) and to make vowel segmentation more reliable (in both cases)*. In addition, /l/ is typically velarized in Catalan which causes a different pattern of coarticulation, especially in coda position. For each context, five words were selected, yielding a total of 210 items (7 vowels  $\times$  3 word length/syllable structure conditions  $\times$  2 preceding consonantal contexts  $\times$  5 words). The complete list can be found in Appendix A. The low/high quality of mid vowels was determined by the first author, who is a native speaker of the variety of Catalan under study, although she is not from the same geographical region as our speakers. We do not expect that all speakers should produce all words according to the qualities assigned to these stimuli: the author's intuitions serve as a necessary standard against which other speakers' judgments are compared. We consulted two dictionaries (Alcover & Moll, 1978; Bruguera i Talleda, 2004), and our intuitions did not always match the transcriptions in the dictionaries, but there was not complete agreement between the two dictionaries either. Furthermore, the dictionaries did not include conjugated verb forms, which made it impossible to use the transcriptions of one of these published sources as our "standard".

Three minimal pairs involving the mid vowels (two with front mid vowels and one with back mid vowels, all disyllabic words) were included because they met all other requirements. While the main goal of the article is to explore the realization of the mid vowels, including the whole vowel system allows us to understand the mid vowels' placement in the vowel space with respect to /i a u/, and to evaluate acoustics across the vowel system.

To be able to meet the requirements in terms of phonetic context, syllable structure, and word length, it was decided to include words belonging to different word classes: mostly nouns and verb forms, but also a few adjectives and function words. However, we did not select these words based on their class, or on expectations that they might show more or less variation in the realization of mid vowels. For instance, in the introduction, we noted that stress-final infinitives do not provide information on the underlying quality of the mid vowel in the stem, and that it has been argued that this causes confusion for the speaker. Yet, nobody has actually tested the possible consequence that verb forms may be less stable than, e.g., nouns. In fact (and with the caveat that our data were not designed to test this claim), to foreshadow one finding of this study, our data (Appendix C) do not seem to corroborate the hypothesis that verbal forms that alternate with a reduced vowel in the infinitive are less stable. Neologisms, borrowings, or learned words were not included, given that their phonological behavior departs from native vocabulary in some respects (Cabr , 2009).

Target words were embedded in carrier phrases. To divert speakers' attention from the target words, the carrier phrases were varied. There were a total of five carrier phrases, with the same number of syllables before and after the target words and the same distribution of stressed and unstressed syllables (see Appendix B). Throughout the experiment and across speakers, each word appeared in the same carrier phrase.

### *3.3 Data collection*

Participants were recorded in a sound-attenuated booth in the Speech Analysis Unit at the Universitat Rovira i Virgili. They wore a head-mounted condenser Audix HT5 Microphone. The data were recorded using a Fostex DC-R302 3-Channel Audio Mixer and Stereo Recorder with a sampling rate of 44.1 kHz.

The data were collected by the first author, a Catalan-dominant Catalan-Spanish bilingual. All communication with the participants was carried out in Catalan. At the beginning of the recording session, after signing the consent form, participants answered some general questions to adapt to being in the sound-treated booth. After the questions, the stimuli were presented on a laptop screen (one sentence at a time) situated outside the booth and visible through a window. Each sentence was recorded three times, presented in three separate blocks (with different random orders). The pace was controlled by the researcher, who was outside the sound booth.

Following the recording session, participants completed a questionnaire about speaker intuitions. The instructions, given in Catalan, asked participants to indicate whether the stressed vowel in each word was a low mid or high mid vowel in their pronunciation. Apart from two minimal-pair target words, none had orthographic stress marks, so that participants could not rely on orthography to make their decisions. There were three possible answers: “low mid” (*oberta*, lit. ‘open’), “high mid” (*tancada*, lit. ‘closed’), and “I don’t know”. Educated speakers of Catalan are familiar with these terms to distinguish between the two sets of mid vowels. In fact, having to classify words depending on whether they have low mid or high mid vowels is a common task that students complete in both primary and secondary school. Basic knowledge of the phonological system of Catalan (including knowledge of the phoneme inventory and phonetic transcription conventions) is tested in the university access exam that all our speakers must have completed, given their status as university students. Participants were asked to keep track of any words they were not familiar with and write them in the comments section at the end of the survey.

The experiment took approximately 45 minutes to complete, and participants were compensated with 10€ for their time.

### 3.4 Data Analysis

A total of 8,610 vowels were collected ( $210 \times 3$  repetitions  $\times$  14 speakers = 8,820; one repetition of one speaker, f12, was lost). For each token, vowel onset and offset were manually segmented. Eight tokens were discarded due to vowel elision or data acquisition problems, yielding 8,602 tokens for data analysis. F1 and F2 were extracted at vowel midpoint using the Burg algorithm in Praat (Boersma & Weenink, 2015) with default settings for female speakers (formant ceiling: 5,500 Hz; number of formants to be extracted: 5). Formant values were checked for tracking errors and manually corrected.

Initial examination of the data in terms of phonetic context revealed that vowels showed coarticulation with flanking consonants. In particular, postlabial vowels had lower F2 than those following dentoalveolar consonants. However, there were no systematic patterns of coarticulation in F1, which is the relevant dimension for vowel height. Therefore, we minimize further discussion of the effects of consonantal context.

For each speaker, we created an integrated data set that included mid vowel judgments and acoustic data, so that F1 and F2 values could be considered in light of either experimenter-assigned mid vowel quality (henceforth *ExperimenterV*) or the speaker’s own judgment of vowel height for a particular token (*SpeakerV*). With this technique it is possible to evaluate (a) the degree to which speakers agree among themselves; (b) whether the speaker’s judgment matches her acoustic realization of a token; and (c) whether a speaker is consistent across all three repetitions of a given word. To test (a), we use the Fleiss’ Kappa test (Fleiss, 1971), carried out using the *kappam.fleiss()* function in the *irr* package (Gamer, Lemon, Fellows, & Singh, 2012) for R (R Core Team, 2000).

In the results that follow, speaker-specific plots and certain measurements use raw F1 and F2 values; however, some portions of the analysis rely on normalized data. Data were normalized using the Lobanov z-score transformation (Lobanov, 1971), which results in speaker-intrinsic, vowel-extrinsic, and formant-intrinsic normalization shown to preserve vowel-specific information, as well as regional or sociophonetic characteristics, while minimizing variation due to anatomical differences from e.g. body size or gender (Adank, Smits, & Van Hout, 2004; Chen, 2008). In this procedure, mean values and standard deviations of F1 and F2 are calculated for each formant (*F*) separately, for speaker (*i*), using data from all vowels. The difference is calculated between the F1 or F2 value of each token (*n*) and the appropriate formant-specific mean; and the resulting value is divided by the formant-specific standard deviation. Once applied to all tokens, the output is centered around a mean of (0, 0), where for example a high front /i/ has negative values for z-scored F1 and positive values for F2, while an /a/ has an F2 near zero and a positive F1. The formula for this transformation is given in (1).

$$(1) z_{ni} = \frac{(F_{ni} - \mu(F_i))}{\sigma(F_i)}$$

We tested the phonetic contrast between high and low mid vowels across speakers using a mixed effects model, which was run in R using the *lme4* package (Bates & Maechler, 2009) with the *lmerTest* package (Kuznetsova, Brockhoff, & Christensen, 2013) to calculate *p*-values. The model was run separately on normalized F1 and F2 data from mid vowels only. Its fixed effects were vowel pair (front: /e ε/; back: /o ɔ/), height (high: /e o/; low: /ε ɔ/), context (dental vs. labial), syllable count, syllable type, and repetition (out of three). Mid vowel height was classified by *SpeakerV*. All possible two-way interactions between pair, height, and context were also included. Random intercepts were fitted for each word, and random intercepts were fitted for each subject, with by-subject random slopes for repetition number.

To test whether individual speakers maintain two separate, distinct vowels for each vowel pair, we calculated Pillai scores for each speaker’s mid vowel contrasts. Pillai scores are an output of a multivariate analysis of variance, which are used to indicate the degree of merger or distance between two clouds of data points (Amengual & Chamorro, 2015; Hall-Lew, 2010; Hay, Warren, & Drager, 2006). Pillai scores for each speaker were based on normalized data, for /e ε/ and /o ɔ/ separately, where vowel height was determined by *SpeakerV*. Factors included in the multivariate analysis were vowel type (high mid or low mid), consonantal context (dental or labial), syllable structure (CV or CVC), and syllable count (2 or 3). Given this range of variables, a high Pillai score (with an accompanying small *p*-value) indicates a significant distance, or lack of merger, between the two vowel types; a small or insignificant Pillai score suggests a merger

(Amengual & Chamorro, 2015; Sloos, 2013). We also determined the extent to which Pillai scores correlate with BLP scores using Pearson’s product-moment correlation tests.

Vowel duration is not analyzed. In Nadeu (2014), the duration between high mid and low mid vowels was shown to be around 10 ms, while the F1 difference was approximately 200 Hz. Thus, F1 seems to be a very robust correlate of the mid vowel contrasts, whereas the length difference may be barely noticeable or even perceivable.

We use the standard deviations of F1 (raw data) to quantify the amount of overlap among vowels in each speaker’s data. More specifically, we calculate the V1/V2\_overlap measurement in Fougerson & Audibert (2011), shown in (2). For a contiguous vowel pair (e.g. /i/ and /e/), the standard deviation of the higher vowel (with lower F1) is added to its mean, whereas for the lower vowel the standard deviation is subtracted from the mean. This produces the upper limit of F1 standard deviation for the higher vowel, and the lower limit of F1 standard deviation for the lower vowel. Finally, the latter is subtracted from the former. A negative value indicates no overlap (that is, a positive distance in Hz) between the two vowels, whereas a positive value expresses the amount of overlap in Hz between the two standard deviations. We computed overlap measurements (F1 only) for the following vowel pairs: /i e/ (front high vs. high mid), /e ε/ (front high mid vs. low mid), /ε a/ (front low mid vs. low); and /u o/ (back high vs. high mid), /o ɔ/ (back high mid vs. low mid), /ɔ a/ (back low mid vs. low).

$$(2) i/e \text{ overlap} = (\mu(F1_i) + \sigma(F1_i)) - (\mu(F1_e) - \sigma(F1_e))$$

Correlations between BLP scores and overlap measurements were run. In addition, an ANOVA comparing the overlap between high mid vs. low mid pairs (averaged across front and back vowels) to the other pairs was carried out. Post-hoc t-tests were used to determine whether the amount of overlap between high mid vs. low mid vowels was larger than that between high and high mid vowels and that between low and low mid vowels. For the t-tests, *p*-values over .025 are considered significant (after Bonferroni correction for the two t-tests).

A k-means clustering analysis was applied to the acoustic data for each speaker’s mid vowels. K-means clustering determines which data points (i.e. realizations of specific lexical items) group together in the high mid vowel space, and which group together as low mid vowels. This technique, applied simultaneously to F1 and F2 data in the two-dimensional vowel space, partitions the data points in order to minimize the sum of squares distance from the points to their assigned cluster. The analysis was applied separately to data from front and back mid vowels, to avoid unintentional clustering across the vowel space. It was applied to speaker-normalized F1 and F2 data, and was conducted in R using the *kmeans()* function, applying the Hartigan and Wong (1979) algorithm with two clusters per speaker (per front/back category), and 100 random starts. The output of the k-means clustering is a list of which words, or data points, belong to each of four mid vowel clusters, which can be compared with vowel height according to ExperimenterV and SpeakerV. We explore questions (b) and (c) above using the k-means clustering data.

## 4. Results

### 4.1 Speaker intuitions vs. experimenter-assigned vowel height

Before showing plots of the vowel space, in this section we quantify how speakers' intuitions about the height of mid vowels in the target words diverge from vowel height assigned by the experimenter (SpeakerV and ExperimenterV, respectively). Speaker-specific, word-specific judgments are listed in Appendix C, while rates of agreement with ExperimenterV, per mid vowel category but averaged across lexical items, are summarized in Table 1. Speakers (considered as a group) agree with ExperimenterV in roughly 80% of words across all mid vowels. The details of these agreements are explored in section 4.2.

**Table 1** *Rates of agreement between speakers' mid vowel judgments and experimenter-assigned vowel height (shown in slashes), averaged across words and speakers.*

Front vowels	Mean agreement (SD)	Back vowels	Mean agreement (SD)
/e/	80.5% (25.6%)	/o/	80.2% (21.8%)
/ɛ/	80% (21.8%)	/ɔ/	79.8% (11.4%)

Table 2 focuses on individual data, summarizing how each speaker's mid vowel height judgments differ from those assigned by the experimenter. Disagreements occur roughly equally often across front vowel words (150 cases; mean = 10.7) and back vowel words (158 cases; mean = 11.3). Some speakers (f01, f02) disagree only a few times with the experimenter's assignment; others (f10, f11) disagree in nearly half of cases (out of 60 words total per front/back). The fact that disagreements do occur sets the stage for comparing vowels in the acoustic space according to ExperimenterV vs. SpeakerV. The fact that speakers disagree at different rates, on the other hand, indicates that if speakers are good judges of their own pronunciation, then even within a relatively uniform dialect area, the height of Catalan mid vowels is not fixed in all lexical items. This is also quite clear from the results in Appendix C.

**Table 2** *Number of disagreements (percentages in parentheses) between experimenter-assigned mid vowel height in target words, and speaker judgments of mid vowel height, divided by target vowel backness. The stimuli include 60 front mid vowels and 60 back mid vowels.*

Speaker	Front vowel disagreements	Back vowel disagreements
f01	6 (10%)	5 (8.3%)
f02	6 (10%)	2 (3.3%)
f03	11 (18.3%)	4 (6.7%)
f04	8 (13.3%)	5 (8.3%)
f05	6 (10%)	6 (10%)
f06	8 (13.3%)	4 (6.7%)
f07	13 (21.7%)	14 (23.3%)
f08	11 (18.3%)	4 (6.7%)
f09	19 (31.7%)	25 (41.7%)
f10	24 (40%)	28 (46.7%)
f11	20 (33.3%)	26 (43.3%)
f12	6 (10%)	6 (10%)
f13	4 (6.7%)	27 (45%)
f14	8 (13.3%)	2 (3.3%)
Total	150 (17.9%)	158 (18.8%)

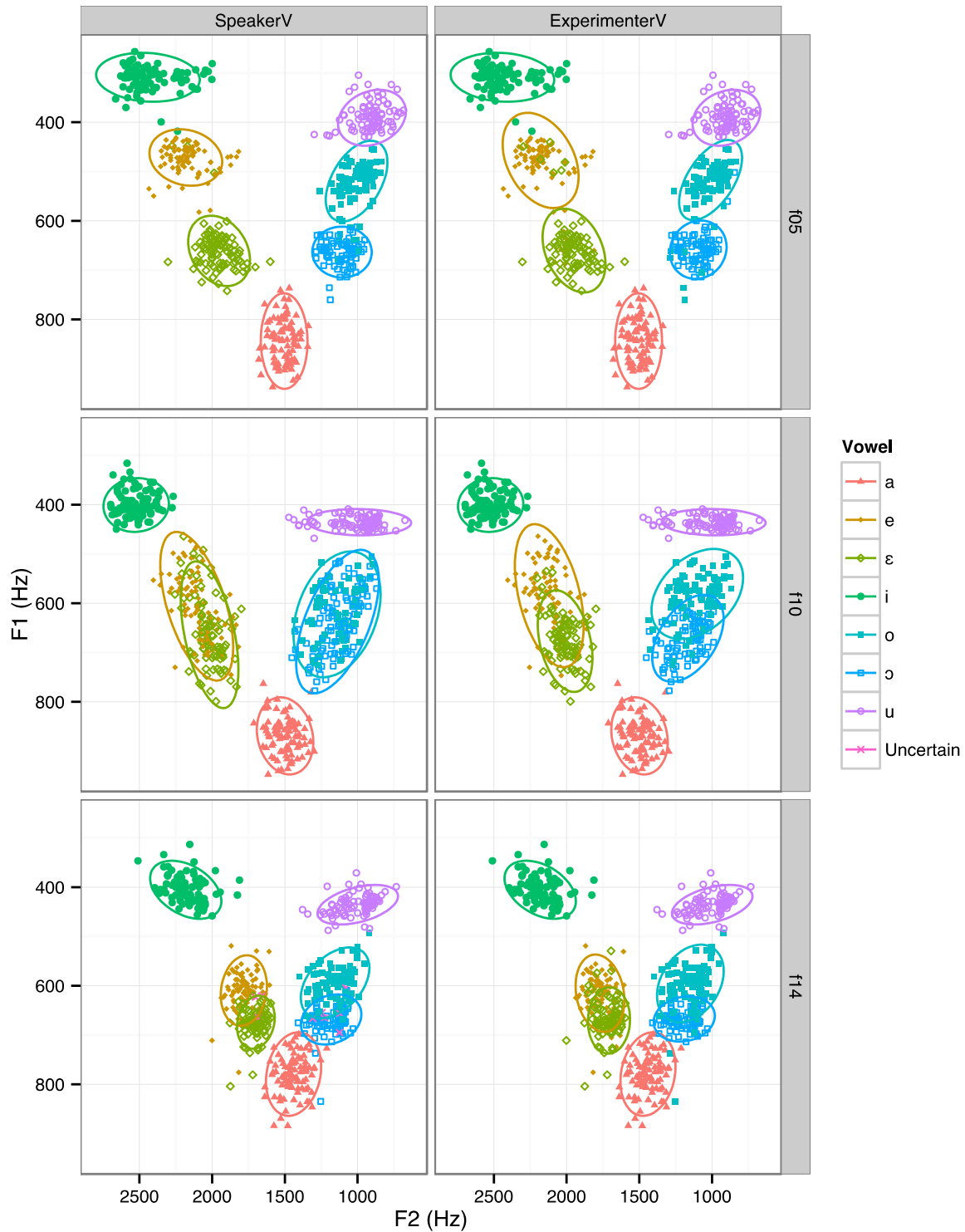
#### 4.2 Interspeaker agreement rates

The speakers who participated in this study all come from the same Catalan dialect area and from a relatively small geographical area, so we expect them to behave quite similarly with regard to the mid vowels, if indeed the front and back contrasts are robust for signaling lexical meaning. Therefore, more interesting than the amount of agreement between the experimenters' intuitions and those of the participants is the agreement among our participants. Out of the 120 words with mid vowels, there are only 17 for which all 14 speakers agree. We tested the degree of agreement among participants with a Fleiss' Kappa test. For each speaker, each word was coded as "high", "low", or "I don't know" (meaning they did not know the vowel height, but knew the word), depending on their answers in the survey. The kappa coefficient returned was 0.50 (0.53 including only the 106 words that all speakers were familiar with), which is interpreted as a moderate agreement (based on the Landis scale, Landis & Koch, 1977). These results indicate that speakers agree at a rate greater than chance. Despite not always assigning the same vowel to the words they were presented with, there is some agreement among them. However, taking into account that these vowels form phonological contrasts, one might expect very high agreement, rather than simply moderate.

When Experimenterv is added to the data as another independent rater, the kappa index is virtually the same (0.51; 0.55 including only the 106 words that all speakers were familiar with). Therefore, there does not seem to be more agreement among our participants than among our participants and Experimenterv. Observation of the judgments in Appendix C does show that there may be some differences between the experimenters and the participants that are due to intradialect variation: the word *orca* 'orca' was judged to have /ɔ/ by 12 speakers (the other two answered "I don't know"), and *setze* 'sixteen' was assigned /e/ by 13 speakers, disagreeing with Experimenterv. However, cases of disagreement among participants are attributed to individual variation.

#### 4.3 The Catalan vowel space

Figure 2 shows data from three speakers who exhibit different patterns in their realization of mid vowels. For each speaker two plots are presented: on the left, the mid vowels are coded by SpeakerV; on the right, by Experimenterv. Plots for all speakers (SpeakerV only) are given in Appendix D. The data points represented are the same. Differences between the SpeakerV and Experimenterv plots are in the labels assigned to mid vowel data points.



**Figure 2** Vowel spaces of three Catalan speakers: f05, f10, f14. For each speaker, the data points represented are the same in the two plots. However, mid vowel height is determined on the left by the speaker's judgment (SpeakerV) and on the right by the experimental standard (ExperimenterV). Ellipses are based on 95% confidence intervals and, for the SpeakerV plots, they exclude any "Uncertain" tokens. Each point represents one repetition of a target vowel.



Despite the vowel overlap seen in some speakers, the high and low mid vowels remain phonetically distinct, a result confirmed by mixed-effects modeling (see Appendix E). The model for F1 shows a significant effect of vowel height ( $p < 0.001$ ). The only other significant factors in the model are vowel pair ( $p < 0.05$ ) and the interaction between pair and height ( $p < 0.05$ ), meaning that the distribution of F1 values varies across front and back vowels. The model of F2 shows a wider array of significant factors, consistent with the coarticulatory effect of consonantal context on backness; but vowel height is not a significant predictor of F2 ( $p = 0.356$ ).

Pillai scores (see 3.4) confirm the distinctness of mid vowels for all speakers. Table 3 shows that for all speakers, the Pillai score is significant for front vowels ( $p < 0.001$ ). Among back vowels, all speakers but two have significant Pillai scores. Although (nearly) all scores are highly significant, the values themselves vary widely, from less than 0.05 to more than 0.80. This indicates wide interspeaker variation in the distance between mid vowel pairs, where larger values indicate greater separation. Table 3 also includes speakers' BLP scores, where a high score indicates greater Catalan dominance. There is no correlation between BLP score and Pillai score across speakers, either among front vowels ( $r = -0.151$ ,  $p = 0.61$ ) or back vowels ( $r = -0.119$ ,  $p = 0.69$ ).

**Table 3** *Individual speakers' BLP scores, indicating level of Catalan dominance, and results of the Pillai test, indicating distance between mid vowels, for front and back pairs, separately.*

Speaker	BLP score	Front Pillai Value	$p$	Back Pillai Value	$p$
f01	96.36	0.50	< 0.001	0.57	< 0.001
f02	75.47	0.80	< 0.001	0.79	< 0.001
f03	86.46	0.76	< 0.001	0.87	< 0.001
f04	105.26	0.81	< 0.001	0.84	< 0.001
f05	88.37	0.86	< 0.001	0.81	< 0.001
f06	93.09	0.64	< 0.001	0.55	< 0.001
f07	87.37	0.50	< 0.001	0.28	< 0.001
f08	86.55	0.63	< 0.001	0.60	< 0.001
f09	117.70	0.28	< 0.001	0.14	< 0.001
f10	75.20	0.17	< 0.001	0.03	n.s.
f11	94.54	0.26	< 0.001	0.03	n.s.
f12	89.91	0.71	< 0.001	0.58	< 0.001
f13	85.91	0.58	< 0.001	0.11	< 0.001
f14	68.39	0.60	< 0.001	0.60	< 0.001

We hypothesize that, where speakers' judgments of mid vowel height diverge from ExperimenterV, the individual speaker's judgment will be a better match for the acoustic realization. This is true for two speakers in Figure 2, f05 and f14. These two speakers largely agree with ExperimenterV; nevertheless, they are better judges of their own pronunciation: for speaker f05, several tokens labeled /e/ by ExperimenterV are realized in the [ɛ] cloud, and vice versa; the same occurs among the back vowels. In the plots for f14, the mid vowel clouds are more distinct in height in the SpeakerV plot than the ExperimenterV plot, despite some overlap in both plots. The case of f10, however, is markedly different: in her SpeakerV plot, both front mid vowel ellipses are realized in the same acoustic space, and likewise for the back vowels. When her vowels are coded by ExperimenterV, on the other hand, vowels labeled /e/ are

generally realized higher in the vowel space than those labeled / $\varepsilon$ /, and likewise for / $o$   $\varnothing$ /. It appears that speaker f10 does not have accurate intuitions of her own mid vowel height; but she does produce a mid vowel cloud that occupies the high and low mid vowel regions. In fact, most speakers seem to be quite good judges of their own realizations, but some (f09, f10, f11, f13) are not. As shown in Table 2, these four speakers disagree most frequently with ExperimenterV. For these speakers, the ExperimenterV labels coincide better with their acoustic realizations than their own judgments do. This finding extends to Pillai scores: speakers f10 and f11 had non-significant Pillai scores for back vowels when SpeakerV was used to calculate them (see Table 3), but the scores were significant when ExperimenterV labels were used (f10: Pillai 0.55,  $p < 0.001$ ; f11: Pillai 0.69,  $p < 0.001$ ).

Based on the vowel plots in Appendix D, speakers were classified according to the degree of phonetic overlap among different vowel categories. Some speakers (e.g. f05 in Figure 2) have seven acoustically distinct vowels, whose ellipses do not overlap. Others have varying degrees of overlap, particularly among the mid vowels: they may have slight overlap (f14); moderate overlap (see f08 in Appendix D); or complete overlap (f10). Table 4 combines this judgment of vowel space overlap with calculations of the amount of F1 overlap between each speaker's front and back mid vowels (see Section 3.4). Recall that a negative value indicates separation between the vowels, and a positive value shows the extent of overlap in Hz between the two vowels.

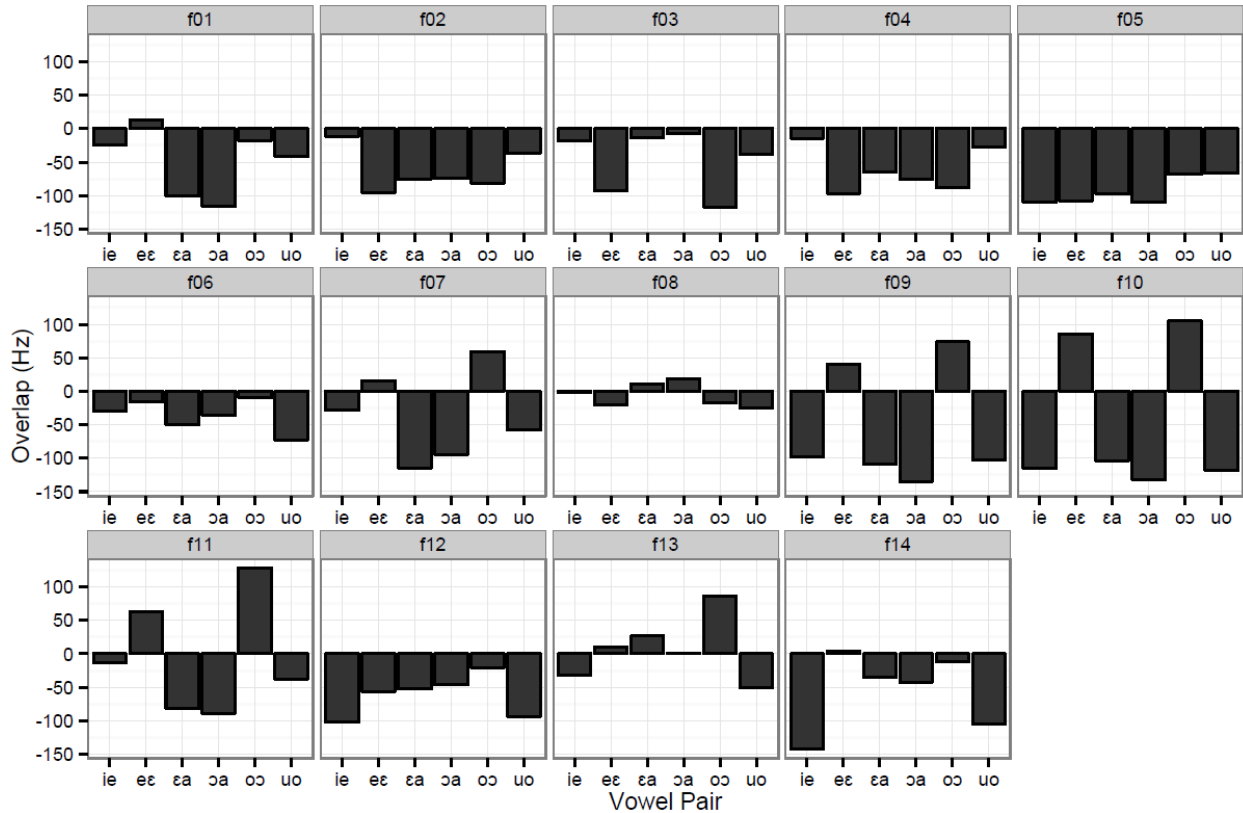
**Table 4** *Type of vowel-space overlap, given mid vowel heights assigned by speaker; and the amount of F1 overlap between high and low mid vowels in Hz.*

Speaker	Vowel space overlap	Front F1 overlap	Back F1 overlap
f01	Slight overlap	12.89	-17.59
f02	Distinct	-95.25	-81.23
f03	Distinct	-92.29	-116.80
f04	Distinct	-96.57	-87.56
f05	Distinct	-107.65	-68.31
f06	Moderate overlap (back > front)	-15.65	-9.76
f07	Major overlap	15.35	59.23
f08	Moderate overlap	-20.29	-17.04
f09	Complete overlap	40.19	73.68
f10	Complete overlap	85.86	105.07
f11	Major overlap (back > front)	62.50	128.79
f12	Distinct	-56.22	-20.44
f13	Major overlap (back > front)	9.70	86.82
f14	Slight overlap	3.86	-11.42

In Table 4 we observe a good correspondence between the amount of mid vowel overlap according to visual classification, and the amount of F1 overlap according to Fougeron & Audibert's (2011) formula. Speakers who have visually distinct mid vowels (f02, f03, f04, f05, f12) have the lowest (always negative) values by this measure, in the front and back of the vowel space, though the front vowels are typically more distinct (e.g. f05: -107.65 Hz for front vowels, but -68.31 Hz for back vowels). Speakers with slight or moderate visual overlap among the mid vowels have positive values by both measures, while speakers with major or complete overlap (f09, f10, f11) have the highest positive values.

Variation in the realization of mid vowels is, thus, not restricted only to the lexical level: speakers do not differ solely in the composition of their sets of high and low mid vowel words. Rather, we find substantial variation in the phonetic implementation of these contrasts, both for the front and back series. We tested whether mid vowel overlap correlated with speakers' dominance in Catalan: Pearson's product-moment correlations between BLP scores and front vowel overlap ( $r = 0.03$ ,  $p = 0.92$ ) and between BLP scores and back vowel overlap ( $r = 0.09$ ,  $p = 0.76$ ) were non-significant. These non-significant relationships echo the lack of correlation between Pillai score and BLP score (Table 3) and are in line with the findings in Amengual (2014), who showed that Euclidean distances between /e/ and /ɛ/ and between /o/ and /ɔ/ did not correlate with BLP score for Catalan dominant speakers (although the correlations were significant for that study's Spanish-dominant group). We thus find no evidence that variation in Catalan dominance among these speakers is linked to different phonetic implementations of the mid vowels.

To evaluate the hypothesis that vowels are equidistant and equally overlapping in the vowel space (see section 2), Figure 3 displays F1 overlap between adjacent vowels for each speaker, and reveals that the acoustic distance between mid vowels tends to be, for most speakers, smaller than the distance between other contiguous vowel pairs.



**Figure 3** *Overlap measurement (in Hz) between contiguous vowel pairs by speaker. Negative values indicate no overlap.*

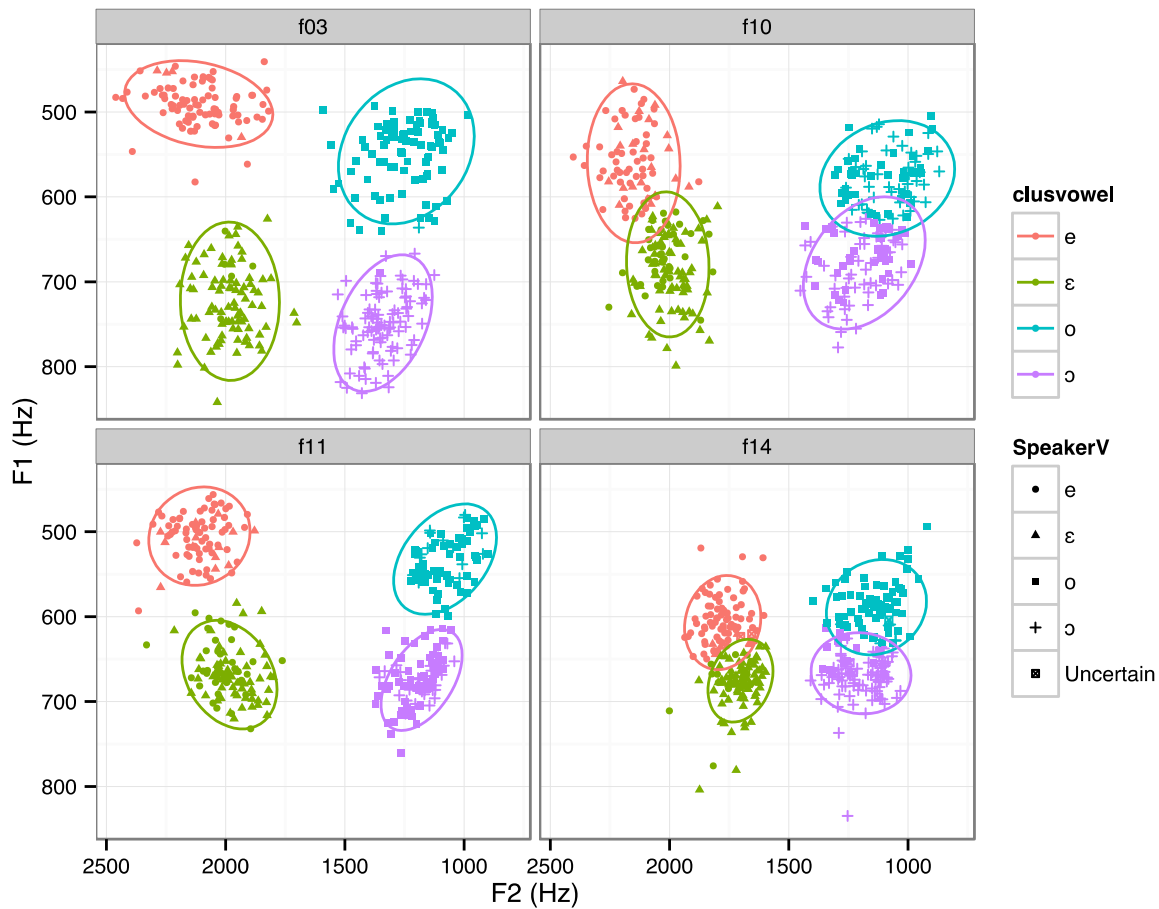
The vowel pairs that show positive acoustic overlap more often are also mid vowel pairs (speakers f01, f07, f09, f10, f11, f13, f14). An ANOVA with overlap as dependent variable and contiguous vowel pair (classified into high vs. high mid; high mid vs. low mid; and low mid vs.

low, see Section 3.4) as a within-subjects factor returned a significant effect of vowel pair ( $F(2, 26) = 4.19, p < 0.05$ ). Post-hoc t-tests were run to explore the vowel pair effect. The high mid vs. low mid comparisons ( $M = -8.22$  Hz,  $SD = 69.85$ ) have higher overlap measurements than low mid vs. low ( $M = -64.24$  Hz,  $SD = 46.64$ ;  $t(47.08) = 3.53, p < 0.001$ ) and high vs. high mid ( $M = -57.60$  Hz,  $SD = 40.30$ ;  $t(43.18) = 3.24, p < 0.01$ ) vowel pairs. These findings indicate that, for most speakers, stressed vowels are not equidistant in the vowel space. Rather, the high mid and low mid vowels are closer in F1 to one another than to other vowels in the system.

#### 4.4 *K-means clustering analysis*

We next address how accurate and consistent speakers were in their judgments and productions. A k-means analysis clusters the data points according to their location in the acoustic vowel space (see section 3.4). The k-means output was used to plot individual vowel tokens according to both the cluster assigned by the k-means algorithm and the speaker's judgment. Several examples of the results appear in Figure 4, where color and ellipse indicate the k-means cluster and shape indicates SpeakerV. While each speaker performs slightly differently, there are patterns in behavior with regard to two elements: whether the speaker realizes high vs. low mid vowels as distinct clouds (f03, f11) or as a contiguous cloud in the vowel space (f14, f10); and whether the speaker is a good judge of her own pronunciation (f03, f14) or not (f11, f10). Out of 14 speakers, six were uncertain about at least one judgment, exemplified here by f14, for whom sparse symbol overlap in fact suggests that she is a good judge of vowel height.

Speakers appear to be somewhat better judges of front vowel height than back vowel height, despite the fact that back and front mid vowels are quite symmetrical in vowel height. This is different from a recent study of Italian (Renwick & Ladd, in preparation), which finds that because of regional differences in phonological conditioning of mid vowels, some speakers are better judges of back vowel height than front vowel height, but that front vowels are generally more distinct in acoustic height than their back vowel counterparts (although conversely speakers have greater confidence in their intuitions about front vowels than back vowels; Ladd, 2006).



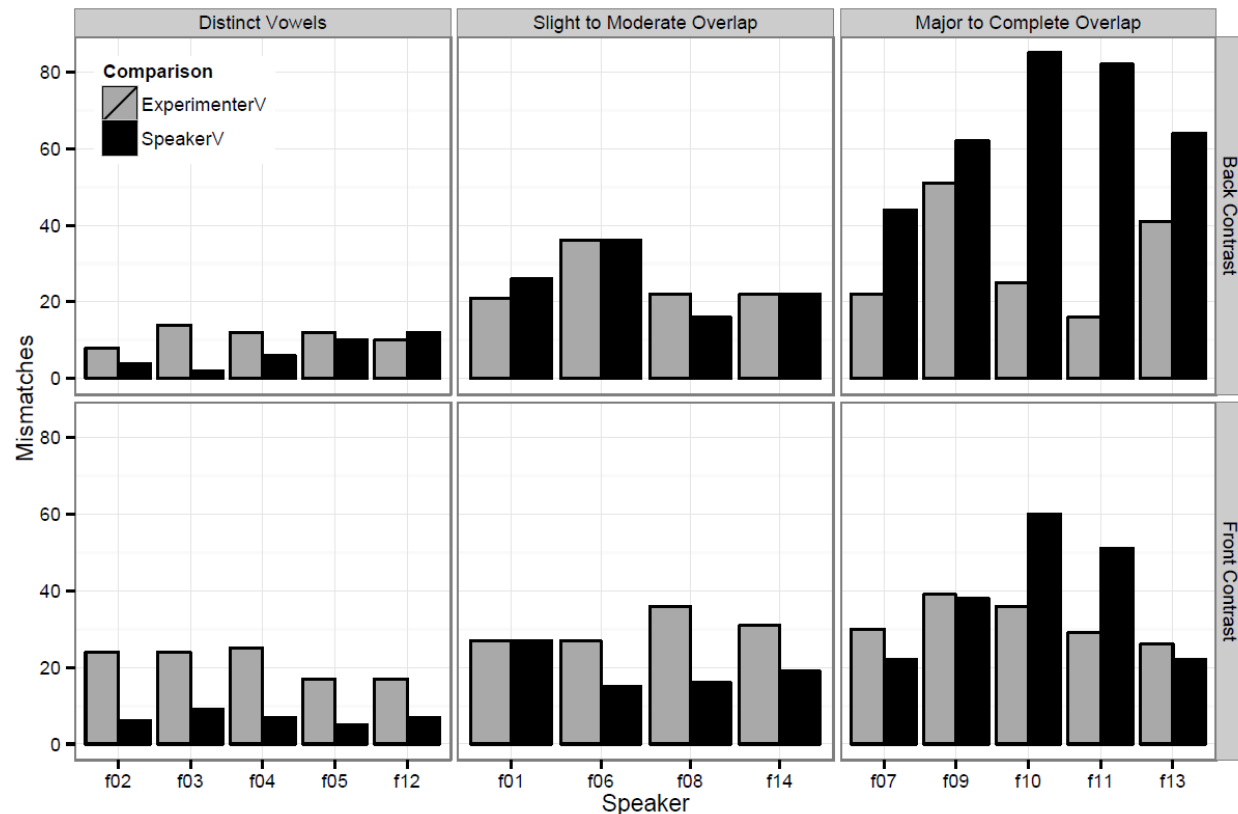
**Figure 4** Mid vowels for four Catalan speakers, labeled by k-means cluster (*clusvowel*; color) and speaker vowel height judgment (*SpeakerV*; shape). 95% confidence ellipses are based on k-means cluster assignment.

#### 4.5 Misclassifications

We use the output of the k-means algorithm to evaluate whether participants are accurate at the task of classifying their own vowels according to height; that is, whether vowels that speakers judge to be high mid in their own speech indeed appear in the higher of two mid vowel clouds (and vice versa for low mids). When the cluster assigned to a specific token by k-means does not match the speaker’s vowel height judgment, a mismatch occurs between the judgment and the acoustics. We also compare the k-means cluster against vowel height as assigned by the experimenter. This section describes how often mismatches occur for each speaker, between the acoustics (k-means clustering) and *SpeakerV* or *ExperimenterV*. We assume that mismatches in the first case reveal how accurate a speaker is at reflecting on her own pronunciation, while the second case highlights differences in intuition between the experimenter and the speaker. If speakers’ productions correspond better to their own judgments than to *ExperimenterV*, then we expect fewer mismatches when speaker judgment is compared to k-means clustering results.

We first discuss cases of mismatch between k-means cluster and speaker intuition, which are plotted in black in Figure 5. To avoid inducing mismatches when a speaker indicated uncertainty, all words whose vowel height a particular speaker could not identify were excluded from this

comparison. Additionally, we excluded from the analysis all tokens for which there was both a mismatch between SpeakerV and k-means vowel cluster, and the speaker was unfamiliar with the word, as they may have resulted from guessing rather than intuition.



**Figure 5** *Number of mismatches between judged vowel height and k-means cluster. Black bars: number of mismatches between SpeakerV and k-means cluster. Gray bars: number of mismatches between ExperimenterV and k-means cluster. Speakers are divided by the amount of acoustic overlap present in their mid vowels (see Table 4).*

Using k-means clustering to divide high and low mid vowels, each speaker has at least two acoustics-judgment mismatches among both front and back vowels. The results in Figure 5 pool across all vowel tokens for each speaker, e.g. up to three repetitions of each target word. If a speaker produces a word with the same height across all repetitions, but reports the opposite height judgment, then that word is counted thrice. We address the issue of consistency across repetitions in section 4.6. As Figure 5 shows, while some speakers have more front mismatches than back mismatches (f02, f03), the majority of speakers have more back mismatches (f06, f09, f11, f13, etc.). This may indicate that the front mid vowels are more acoustically distinct for speakers than the back vowels, and may correlate with speakers' strength of intuition of vowel height (this has not been tested). Some speakers, especially f09, f10, f11, and f13, have between 60 and 80 mismatches among back vowels (out of 360 tokens, but 240 for speaker f12).

The results of a comparison between ExperimenterV and k-means are also given in Figure 5 (gray bars). Certain data points were excluded, with the same methods used for the SpeakerV-acoustics mismatches. Using the k-means clusters to divide high from low mid vowels, each

speaker has at least eight mismatches between acoustics and ExperimenterV among both front and back vowels. With respect to the mismatches between SpeakerV and acoustics (black bars), the gray bars are more uniform, across speakers but also across front and back, for each speaker. For some speakers (e.g. f02, f03, f04, f05, f12), there are more mismatches between ExperimenterV and acoustics than between SpeakerV and acoustics, indicating that their intuitions correspond better to their phonetic realizations than ExperimenterV does. However, for speakers who had high numbers of mismatches in the first case (f09, f10, f11, f13), the number of mismatches between ExperimenterV and acoustics is much lower. Note that these four speakers had already been identified above as poor judges of their own realizations. As also suggested by their Pillai scores, these speakers' intuitions are sufficiently weak that ExperimenterV is a better predictor of their acoustics. Although this precise comparison has not been carried out for Italian (Renwick & Ladd, in preparation), it is clear in the results of that study, which uses methods similar to ours, that speakers are better judges of vowel height than the prescriptive standard to which judgments are compared; thus, it is surprising that for Catalan the same does not hold across all speakers.

A reviewer notes that those speakers identified as poor judges (f09, f10, f11, f13) may be influenced by Spanish. However, their responses to the BLP questionnaire do not support this interpretation. Like the rest of our speakers, all of them were exposed to Catalan from birth and they all rated their proficiency higher in Catalan than in Spanish. Speakers f09–f11 reported using only Catalan to communicate with their families and to talk to themselves (f13 uses 10% Spanish, 90% Catalan in these two situations). All of them use Catalan at least 80% of the time at work/university and when speaking with friends.

#### 4.6 Consistency

As described above, the plots in Figure 4 and the counts in Figure 5 include all repetitions of all target vowels in the study; however, each lexical item was only judged once. To fully quantify the ability of speakers to judge mid vowel height in their own speech, we tabulate whether all repetitions of each word are realized in the same vowel cluster, according to the k-means analysis. We calculated, for each speaker, how many times each word was assigned by k-means clustering to a particular vowel cloud. If the word received the same value every time (three times, or twice for f12),<sup>9</sup> that word was “consistent.” If two values were assigned, that word was “inconsistent.” The number of “consistent” words per speaker was tabulated out of the total number of lexical items per front/back category (60), and transformed into a proportion. In Table 5, each speaker's productions are classified according to the number of words that were consistently realized as either high or low mid vowels vs. the number that were realized variably. Words for which the speaker expressed uncertainty or unfamiliarity were not excluded; instead, their count is included in parentheses for each cell.

All 14 speakers realize at least one lexical item (per front/back) inconsistently, i.e. as a high mid in at least one repetition, and a low mid in at least one repetition. Speakers f03 and f05 are the most consistent, while f09 and f14 are the least consistent. Speakers are, overall, slightly more consistent in their realizations of front vowels than back vowels, with 98 total inconsistencies in the former, and 142 in the latter. Interestingly, speakers tended to be consistent in their

---

<sup>9</sup> F12 only has two repetitions of each token.

productions of words for which they expressed uncertainty, or with which they were unfamiliar. For example, speaker f07 was uncertain about eight target words with front vowels, but she produced them all consistently in either the high or low mid vowel cloud. Speakers had greater uncertainty, but also greater consistency, with target front vowels than target back vowels. What is not known is whether speakers find judgments more difficult for words they produce inconsistently.

**Table 5** *Number of target words produced consistently by each speaker, i.e. for which all three repetitions of the lexical item were classified by k-means clustering as having the same mid vowel height. Speaker f12 produced only two repetitions. Parentheses indicate the subset of words in each cell that the speaker was unfamiliar with or whose vowel height she could not judge. No data were excluded.*

Speaker	Front-vowel target words			Back-vowel target words		
	Consistent	Inconsistent	Percentage consistent	Consistent	Inconsistent	Percentage consistent
f01	55 (4)	5	91.7%	45 (1)	15 (2)	75%
f02	56 (1)	4	93.3%	54 (2)	6 (2)	90%
f03	58	2	96.7%	58 (1)	2	96.7%
f04	55 (2)	5	91.7%	55 (1)	5	91.7%
f05	59	1	98.3%	57	3	95%
f06	55 (2)	5	91.7%	44 (2)	16	73.3%
f07	50 (8)	10	83.3%	52 (1)	8	86.7%
f08	52	8 (2)	86.7%	48 (2)	12	80%
f09	47	13	78.3%	43	17	71.7%
f10	50	10	83.3%	47	13	78.3%
f11	54	6	90%	54	6	90%
f12	53 (3)	7	88.3%	54 (3)	6	90%
f13	51	9	85%	44	16	73.3%
f14	47	13 (1)	78.3%	43 (1)	17 (1)	71.7%
Total	742 (20)	98 (3)	88.3%	698 (14)	142 (5)	83.1%

#### 4.7 Minimal pairs

This subsection evaluates how speakers judged, and produced, the three minimal pairs included in the experimental stimuli. These were *néta* ['netə] 'granddaughter' vs. *neta* ['netə] 'clean, f.', *Pere* ['perə] 'Peter' vs. *pera* ['perə] 'pear', and *dóna* ['donə] '(s/he) gives' vs. *dona* ['donə] 'woman'. All three are distinguished orthographically, either by a diacritic (indicating a high mid vowel) or by stronger orthographic distinctions that are neutralized due to vowel reduction (the final vowel in *Pere* vs. *pera*). The results of this evaluation are summarized in Table 6. As shown in the "Judgment" column for each word, speakers largely agreed with experimental vowel height for all three pairs: eleven of the 14 speakers distinguished the minimal pairs in their judgments, by assigning them opposite height values. Speaker f07 assigned high mid status to all but *pera*, while f11 judged both *dóna* and *dona* to have high mid /o/; f13 judged both *néta* and *neta* to have high mid /e/, and f01 was uncertain about vowel height in *pera*.

Additionally, the extent to which speakers' judgments of vowel height in minimal pairs matched their productions was evaluated by comparing judgments with the height assigned by k-means clustering, on a token-by-token basis. The column in Table 6 labeled "# match" shows how many tokens, out of three, were produced in the cluster matching the speaker's judgment. Many



speakers were accurate in their judgment and acoustic distinctions between these minimally distinct words; generally, speakers whose vowel clusters are acoustically distinct were the most successful (f02, f03, f04, f05). Mismatches are more commonly found for speakers who exhibited moderate to complete acoustic overlap, such as f09, f11, f13. Occasionally, speakers were completely wrong, e.g. f01's judgment vs. acoustics for ['netə]. It is worth noting that speakers were most consistent in their judgments for ['netə], but had the fewest number of matches (24/41 possible) for that word, indicating an inconsistency in its production that is unexpected given the interspeaker agreement. Additionally, speakers were most successful at acoustically distinguishing the most orthographically distinct pair, *Pere* (40 matches) vs. *pera* (39 matches). We cannot discard the possibility that the experimental conditions were insufficient for speakers to visually distinguish the minimal pairs that differ in the presence vs. absence of a diacritic.<sup>10</sup> This would explain why they perform so much better with the *Pere-pera* pair, which is distinguished by the initial capital and final vowel.

**Table 6** *Summary of speaker judgments (V) and acoustics-judgment matches for minimal pairs. The form in brackets reflects ExperimenterV, with which other authors agree; “# match” indicates the number of tokens in which the speaker’s judgment and acoustics corresponded to the same vowel height. Assignment of vowel quality in the acoustics was determined via k-means analysis.*

Speaker	['netə]		['netə]		['perə]		['perə]		['donə]		['donə]	
	V	# match	V	# match	V	# match	V	# match	V	# match	V	# match
f01	e	0	ε	3	e	3	?	3 <sup>11</sup>	o	3	ɔ	3
f02	e	3	ε	3	e	3	ε	3	o	3	ɔ	3
f03	e	3	ε	3	e	3	ε	3	o	3	ɔ	3
f04	e	3	ε	3	e	3	ε	3	o	3	ɔ	3
f05	e	3	ε	3	e	3	ε	3	o	3	ɔ	3
f06	e	2	ε	3	e	3	ε	3	o	2	ɔ	3
f07	e	1	e	2	e	3	ε	2	o	2	o	0
f08	e	3	ε	3	e	3	ε	3	o	1	ɔ	3
f09	e	2	ε	3	e	3	ε	3	ɔ	0	o	1
f10	e	1	ε	3	e	3	e	3	o	3	ɔ	3
f11	e	1	ε	3	e	3	ε	3	o	2	o	0
f12 <sup>12</sup>	e	1	ε	2	e	2	ε	2	o	2	ɔ	2
f13	e	1	e	0	e	2	ε	3	o	0	ɔ	3
f14	e	0	ε	3	e	3	ε	2	o	3	ɔ	3
Agreements	14		12		12		14		13		11	
Matches		24		37		40		39		30		33

These results are consistent with the general findings of this study (showing variation across speakers and variable realizations within the same speaker), although both inter- and intraspeaker

<sup>10</sup> Participants were seated in a sound-treated booth facing a glass panel. The computer where the stimuli were presented was situated outside the booth.

<sup>11</sup> Speaker f01 consistently produced this word with [ε] according to the k-means clustering.

<sup>12</sup> Recall that only two repetitions were recorded for speaker f12.

variation seems to be less extensive for the minimal pairs. However, it remains true that certain speakers (e.g. f01, f14 for ['netə] vs. ['netə]; f13 for ['donə]) failed to distinguish these words in production, even when a diacritic was present to indicate (prescriptive) vowel height. Those speakers nonetheless subsequently gave the prescriptively correct judgment for that word. We speculate that these speakers may not have seen the crucial diacritic atop the vowel during recording, and thus this may reflect speaker error rather than authentic acoustics-judgment mismatch.<sup>13</sup> *Néta* and *dóna* are indeed the only stimuli containing such diacritics on the target vowel. Likewise, a potential explanation for f07's across-the-board high mid judgments for these words may be that she did not know how to interpret the diacritics that prescriptively distinguish high from low mid vowels. Despite these factors, it is clear that most speakers do make a distinction in judgment between the members of these minimal pairs, even if they do not always concur with the prescriptive standard.

## 5. Summary and Discussion

### 5.1 Variation in Lexical Distribution

One goal of this paper was to provide evidence for the claim that certain words involving mid vowels do not have a single (standard) pronunciation in Central Catalan. Despite including very few words identified as having variable realizations (Badia i Margarit, 1969a, 1970a; Recasens, 1993), we find substantial variation in our data set. Of the 120 words with mid vowels examined, only 17 were assigned the same vowel quality by all 14 speakers. For all other words, between one and seven speakers disagreed with the rest, leading to only moderate interrater agreement. While this suggests a substantial degree of systematicity, one would expect almost perfect agreement in the distribution of members of a phonological contrast among speakers of the same dialect. Thus, we provide experimental confirmation that goes beyond the available anecdotal evidence to support the claim that the height of mid vowels in certain Catalan words is not fixed.

As discussed in section 3.2, words were selected to comply with certain phonological criteria (phonetic context, word length, stressed syllable structure), which resulted in the inclusion of words that were not matched for other factors that may condition their behavior (for example, word class, lexical frequency, age of acquisition, etc.). Unfortunately, our data do not allow us to examine what conditions the degree of variability exhibited by different words. In fact, target words were not selected because of an a priori expectation that some would show more or less variability than others (with the exception of the minimal pairs, for which we expected little to no variation). Research is under way to determine why and how words vary in the degree of agreement across speakers they elicit, one of the findings of this study.

Despite this variation in the lexical distribution of mid vowels, all of our speakers produced both low and high mid vowel qualities. We can conclude that the language maintains mid vowel contrasts, and so do the speakers who participated in this study. Therefore, the finding that our Catalan-dominant speakers only moderately agree in the lexical distribution of mid vowels is

---

<sup>13</sup> A reviewer notes that, even though the stress mark in these words can only indicate high mid vowel quality, speakers would need declarative knowledge (knowing which of the two members of the minimal pair is spelled with the stress mark) rather than procedural knowledge (knowing how to signal the different vowel qualities) to evaluate the written form of the word. The metalinguistic [process involved in mapping](#) the written form with the correct member of the minimal pair may be hindered, especially during the production test, due to the relatively short time speakers have to start producing the answer.

problematic with regard to phonological contrast. The finding is problematic because even in the speech of one individual, the same phonetic vowel distinction is sometimes meaningful, but sometimes it is not. This extensive variation implies that in everyday communication, speakers encounter words that do not contain the mid vowel that matches their own lexical representation. Yet, presumably, this does not impede successful communication. For example, when encountering the less-likely form ['testə], most speakers (who judged *testa* to have a high mid vowel) must map that form onto the word *testa* /'testə/. This requires speakers to recognize that the high and low mid vowels can, but do not necessarily, contrast meaning. In other words, the same phonetic distance signals meaning in cases of minimal pairs, but must be ignored in a number of words containing mid vowels. This scenario may help account for the amount of variation found in how speakers realized the mid vowel contrasts in production, which we discuss next.

### *5.2 Variation in the phonetic implementation of the contrast*

It was hypothesized that, after controlling for vowel height disagreements, speakers would exhibit a similar phonetic implementation of the contrast. However, the data present three patterns in the realization of the front and back mid vowels: two neatly separated clouds corresponding to the low and high mid vowels, for both the posterior and anterior series; slight to moderate overlap in F1 space; and substantial overlap. Dominance in Catalan did not correlate with Pillai scores or degree of F1 overlap between the low and high mid vowels. At this end of the BLP continuum where Catalan dominance is high, we find wide variation in the separation of the mid vowels that does not correlate with presumed increase in exposure to Spanish, indicating that the strength of the contrast is not solely mitigated by Spanish influence (cf. Recasens, 1993). In addition, the mid vowels are closer or more overlapped with each other in the acoustic vowel space than they are with either high or low vowels. Therefore, the phonetic implementation of these contrasts, like their interchangeability in certain words, is also exceptional compared to other vowel contrasts.

A secondary finding is that, when overlap was observed, it was more common and extensive for the posterior mid vowels. Furthermore, mismatches between labels assigned by the speakers and acoustic realizations were more numerous among back vowels, indicating a clear asymmetry between the treatment of the front mid vs. back mid vowels. Given the complicated evolution of the front mid vowels (see Section 1.2), less robust behavior might be expected for those rather than the back vowels. However, since this is the opposite of what we found, we offer two plausible explanations: one phonological, one phonetic. The phonological explanation is that the back vowel contrast is weaker than the front vowel contrast, leading to greater variation. Alternatively, these front-back asymmetries may reflect universal tendencies in vowel production. de Boer (2011) showed that, crosslinguistically, /u/ tends to have higher F1 than /i/. This entails a more compressed F1 range for the back vowels, which may result in mid back vowels being closer together and exhibiting more overlap. In fact, both of these explanations may be at play. More overlap or a shorter distance between the back vowels may lead to greater speaker uncertainty, hence the higher number of mismatches.

### *5.3 Lexical representations*

The final question concerned whether speakers had clear intuitions of which vowel appeared in a given word. It was expected that speakers would be able to identify the vowel that they produce

in a word and that their realizations of that word would reflect their intuition. The results showed that, whereas most speakers are good judges of their pronunciation, some are not, so much so that the labels assigned by the experimenters matched their acoustic realizations better than their own judgments did. Another surprising finding was that some speakers were not consistent in how they pronounced the very same word across repetitions. That is, it is not only the case that the same word may receive either mid vowel depending on the speaker, but the same speaker may also oscillate between high and low mid realizations. Bosch and Ramon-Casas (2011, p. 521) interpreted differences between Catalan- and Spanish-dominant bilinguals' performance with the mid vowel contrasts as revealing "unstable representations of these vowels in their lexicon" for the latter group. Similarly, Simonet (2014) claimed that these two groups of speakers differ in their phonolexical representations.<sup>14</sup> Our results expand those previous findings by showing that Catalan-dominant speakers alone also differ in their phonolexical representations. In addition, the inconsistency with which certain words are produced suggests that (some of) their phonolexical representations are not very robust.

#### *5.4 General Discussion*

The findings of the current study are consistent with the behavior of mid vowels in other Romance varieties. In several other varieties of Catalan, one or both mid vowel contrasts have been neutralized (Recasens & Espinosa, 2009). In French, the mid vowel contrasts have been reinterpreted as allophonic variation in Southern France (Dufour, Nguyen, & Frauenfelder, 2007), whereas in northern Metropolitan varieties the distinctions are maintained, but they are weakening (Fagyal, Hassa, & Ngom, 2002; Landick, 1995). Turning to Portuguese, Wetzels (1992) argues that, despite being contrastive, mid vowels are neutralized in several contexts due to prosodic, phonotactic, and morphological factors, rendering the contrast "only marginally exploited in B[razilian] P[ortuguese]" (p. 19). Likewise in Standard Italian, few minimal pairs distinguish the mid vowels, and the contrast is treated as marginal (Vincent, 1988). There is great variation in the implementation of the contrast across dialects and regional varieties. In central Italian varieties (e.g. Tuscan, Roman) speakers realize all four mid vowels as distinct, and closely match prescriptive standards; but northern varieties have some phonological conditioning of front vowels, and southern varieties may have heavy conditioning (Avolio, 1995) or entirely lack a mid vowel contrast. When judgments of Italian mid vowel quality are compared to their acoustic realizations, Central speakers tend to be the most accurate, while Northern speakers exhibit some mismatches, and a small sample of Southern speakers indicates a lack of awareness of mid vowel height (Renwick & Ladd, under review).

Thus, across Romance varieties that maintain the mid vowel contrasts, they are markedly different from other vowel contrasts. This difference is characterized by strong variation among and within dialects and individual speakers, the tendency towards merger or allophony, increased phonetic overlap, and low functional load.

Cross-linguistically, it is common to find contrasts of varying strengths: Scobbie & Stuart-Smith (2008, p. 106), acknowledging that all phonemes are not equal, argue that "every language has a rump of potential / actual near-phonemes" (emphasis in the original). In the context of lexical

---

<sup>14</sup> The participants in Simonet (2014) had lower Catalan dominance scores than our participants. It is possible that these differences in BLP scores partially reflect sociolinguistic differences between Catalonia and the Balearic Islands (where Simonet's participants were recruited).

phonology, Goldsmith (1995) describes five types of relationships among sounds, ranging from full contrast to allophones in complementary distribution. These scenarios bracket intermediate cases in which one member of the pair is at least partially conditioned and predictable by phonological rule. The near-complementary relationship between Romanian /i ʌ/ illustrates one such pair; the high central vowel appears under stress, in pre-nasal position and often word-initially, while the mid central vowel is typically word-final and unstressed (Renwick, 2014).

In Catalan, there are cases of completely allophonic relationships among vowels. For instance, the stress-based alternations in Figure 1 occur predictably and almost without exception (see Mascaró, 2002; Nadeu, 2016), to categorically restrict the vowel inventory in unstressed syllables. Our study examines only stressed vowels, where seven contrastive vowels remain; but the strength of some of these contrasts is shown to be weak. We find variation in the assignment of mid vowels to lexical items, across speakers but crucially also within individual speakers' productions. Because this variation exists, we argue that with respect to contrasts like /i/ vs. /a/ or even /i/ vs. /e/, the mid vowel oppositions are functionally less crucial for marking lexical distinctions. We also find evidence of increased phonetic overlap; there are few minimal pairs, suggesting a low functional load; and in some dialects, the mid vowels have merged entirely. In all these senses, the phonemic robustness of Catalan mid vowels is weakened. Specifically, this weakness may reveal erosion of the link between lexical representations, native speaker intuitions, and phonetic production. For Catalan, this claim is in line with Mora et al.'s (2015) hypothesis of the perceptual or phonological weakening of mid vowel contrasts.

Returning to Scobbie & Smith's (2008) premise that not all contrasts are equally contrastive, it is possible to speculate about whether a contrast is "fuzzy" at the individual level, or across members of a language community, or both. Our results indicate that the mid vowel contrasts are phonologically weak both for (some) individuals, for whom phonetic overlap and variability across word repetitions is high, as well as across individuals (at least for certain words). As mentioned before, this entails that the contrasts are not always operative and that speakers must be able to ignore irrelevant variation in some cases.

Acknowledging the "phonological closeness" described for some Romance mid vowel contrasts (Ladd, 2006; Trubetzkoy, 1939) can help us understand how listeners successfully decode lexical meaning, given a production that does not match their own expectations of vowel height. That is, if a listener hears ['petə] and that does not match the listener's intuitions, why is it mapped to /'petə/, and not to /'pitə/, if both are real words? In most instances, speakers encounter words in context, so they can use top-down processes to infer the meaning of the unexpected word. However in the absence of context, we still predict that listeners would associate the unexpected mid vowel with a mid vowel (and not a high or low vowel). First, the overlap data showed that the distance between mid vowels is smaller than the distance between adjacent mid and low or high vowels. We predict that when listeners encounter any unexpected mid vowel, they would tend to classify it as the opposite mid vowel rather than as a point vowel /i a u/. However for unexpected [e] tokens in particular, listeners might tend to classify these as lower [ɛ] because they do not match the precise focal quality required for perception of [i] (Schwartz, Boë, Vallée, & Abry, 1997 and references therein). In addition there are other factors, related to the distribution of mid vowels in the language, which can further explain the interchangeability of mid vowels.

Some borrowings present two phonological adaptations (Cabr , 2009) that involve alternations between high mid (in less nativized loanwords) and low mid vowels (in more nativized adaptations; e.g. *gueto* ['geto] ~ ['getu] 'ghetto'; *p  quer* ['poker] ~ ['p  k  r] 'poker'). In addition, as noted in the introduction, mid vowel height also oscillates in morphologically related words due to lexical and morphological factors (*pes* ['pes] 'weight' vs. *pesa* ['pes ] '(s/he) weighs', *carb * [k  r'  ] 'carbon' vs. *carb  nic* [k  r'  nik] 'carbonic', *conreu* [kun'r  ] 'crop' vs. *conrear* [kunre'a] 'to grow'). The fact that in most words where they appear the high and low mid vowels are spelled using the same orthographic symbol (*e* for the front mid vowels, *o* for the back mid vowels) may also be partially responsible for the equalization of these sounds.

The findings from our production study show, like previous psycholinguistic research, that the mid vowel contrasts are maintained by Catalan-dominant speakers. While previous studies found that Catalan-dominant speakers perceive and produce these contrasts, we noted some patterns in their results that already hinted at the potential lack of robustness or different behavior of these vowels as compared to other phonological contrasts (although these were not examined in the original articles). The extensive individual variation that we obtain in the production of the contrasts is previewed, although not explored, by Bosch and Ramon Casas (2011). Similarly, some studies ran into "problems" with some of the words selected for their studies (Bosch & Ramon Casas, 2011; Sebasti  n-Gall  s et al. 2009; Ramon Casas et al. 2009), indicating, as we demonstrate here, that speakers' phonolexical representations of the same word may differ. Our findings suggest that these contrasts are not only challenging for Spanish-dominant bilinguals, but also for Catalan-dominant ones.

Our results indicate variation at the lexical and phonetic levels, which is consistent with previous findings of perceptual instability (Mora et al., 2011). Anecdotal reports stretching back more than a century indicate variation in the realization of these contrasts at the lexical level. The persistence of this variation could result in not only variability in phonetic implementation, but decreased perceptual sensitivity to the phonological contrast. It is unclear whether the fact that the mid vowel distributions, for two thirds of our speakers, overlap slightly to substantially is the cause or a consequence of weak representations. In an exemplar model of phonology, sublexical units (syllables, individual segments) establish lexical connections, forming exemplar clusters (Bybee, 2010, pp. 20–21). For the mid vowels, because of the conflation of high and low mid vowels as part of the same representational cluster (that is, corresponding to the same lexical item), the clusters mapped to each mid vowel are expected to exhibit more variability and more overlap than, for example, high vowels.

If the mid vowels were strong phonological contrasts, they would be ideal to test questions of bilingualism, given that Spanish has only one vowel where Catalan has two. However, as discussed, the contrast is weak even for Catalan-dominant speakers. Considering learners who are dominant in Spanish, we indeed wonder how they can acquire a contrast that is so variable in reality, when the acquisition of that contrast involves splitting a single phonemic category in the dominant language.

To understand the task facing the (L1 or L2) learner, it would be useful to calculate the functional load (Hall, 2009; Hockett, 1966; Surendran & Niyogi, 2006) of these contrasts vis- -

vis the functional load of other vowel contrasts. However, such calculations require a corpus labeled for mid vowel height by some reliable standard. In Catalan, even different dictionaries vary on their classifications of mid vowels. If an appropriate corpus were available, these data nonetheless present a clear challenge for calculations of functional load. Given a corpus labeled according to some "dictionary standard," we predict mismatches between that standard and speakers' actual productions. On the other hand, if a corpus were labeled according to speakers' actual realizations, then different tokens of the same lexical item would have different heights, potentially even within a single speaker. We hope that future work will be able to address this important question relating to everyday speech communication among Catalan speakers. In particular, it is necessary to survey a wider range of lexical items, with a variety of phonological contexts, from a larger group of speakers, in order to better understand how these contrasts can at once be variable, while remaining a salient aspect of the vowel system.

## 6. Conclusions

Our study is the first to systematically focus on variability in production of the front and back vowels in Catalan-dominant speakers. As we do not assume that a "standard" pronunciation is realistic among speakers of Catalan as a whole or even within a single dialect area, at least for some words (Recasens, 1993), we employ a technique that examines speakers' productions alongside their intuitions. The contributions of this work are both descriptive and theoretical. First, we provide experimental evidence that confirms anecdotal observation and descriptive claims that Catalan mid vowels are somewhat unstable: they trend toward merger for some speakers, or are used interchangeably and inconsistently by others. In addition, this variation occurs among very Catalan-dominant speakers, for whom we would expect the contrasts to be maximally robust. It affects both the lexical distribution of stressed mid vowels and the phonetic implementation of their contrasts, revealing individual speaker differences in both. These findings support Mora et al.'s (2015) phonological weakening hypothesis positing that the Catalan mid vowel contrasts are not robust, and corroborate findings of instability in mid vowel contrasts across Romance varieties.

With regard to theory, our results support the view that not all phonological contrasts are equally robust. In particular, the mid vowel contrasts are less contrastive than others in Catalan because, while they are certainly exploited in some words, such as minimal pairs, they are probably ignored in other cases. This entails that at least certain words with stressed mid vowels have phonolexical representations that are less compact and stable than words with other vowels. Phonological theories of contrast must expand to gain the flexibility required to represent not just phonemes vs. allophones, but also lexical contrasts whose realization may cross phonetic category boundaries even in within a single speaker.

## Acknowledgments

We would like to thank Joaquín Romero for facilitating our data collection at the Universitat Rovira i Virgili (Speech Analysis Unit). We would also like to thank D. R. Ladd, Miquel Simonet, the Associate Editor Marija Tabain, and two anonymous reviewers for their insightful comments and suggestions. Any remaining errors are our own.

## Appendix A. Target words

The following table contains the 210 target words included in the study, classified by identity of the stressed vowel, word length/syllable structure, and preceding consonantal context.

Vowel	'CV.CV	'CVC.CV	CV'CV.CV
/i/	<i>Dentoalveolar</i>		
	tira 'strip'	prisma 'prism'	petita 'small, f.'
	cita 'date'	trista 'sad, f.'	retira 'removes'
	Rita 'Rita'	crispa 'irritates'*	estira 'pulls'*
	frisa 'gets impatient'*	sisme 'earthquake'	recita 'recites'*
	til·la 'lime blossom'	crisma 'understanding'	escrita 'written, f.'
	<i>Labial</i>		
	fita 'milestone'	pista 'hint, clue'	palpita 'beats'*
	pita 'pita'	pizza 'pizza'	aspira 'aspires to'
	pira 'pyre'	pispa 'steals, pop.'*	respira 'breathes'*
	fira 'funfair'	pisca 'nibbles'*	guspira 'spark'
	pila 'battery'	firma 'signature'	confita 'preserves'*
	<i>Dentoalveolar</i>		
	resa 'prays'*	resta 'subtraction'	sencera 'entire, f.'
	rere 'behind'	testa 'head'	hereta 'inherits'
/e/	treta 'removed, f.'	tretze 'thirteen'	barrera 'barrier'
	pressa 'hurry'	cercle 'circle'	pantera 'panther'
	néta 'granddaughter'	presta 'loans'*	enceta 'starts'*
	<i>Labial</i>		
	fera 'beast'	pesta 'plague'	espera 'waits'*
	Pere 'Peter'	pesca 'fishing'	esfera 'sphere'
	peça 'piece'	petja 'gait'	severa 'severe, f.'
	pedra 'stone'	vespa 'wasp'	desfeta 'undone, f.'
	pesa 'weighs'*	festa 'party'	esvera 'alarms'*
	<i>Dentoalveolar</i>		
	cera 'wax'	fresca 'fresh, f.'	mestressa 'owner, f.'
	zeta 'zee'	cresta 'crest'	estreta 'narrow, f.'
	fressa 'buzz'	certa 'true, f.'	empresa 'enterprise'
	neta 'clean, f.'	Cesca 'proper noun'	aterra 'lands'*
	tela 'fabric'	setze 'sixteen'	estesa 'hung out, f.'
	<i>Labial</i>		
	peta 'breaks'*	perdre 'to lose'	cabeça 'bulb'
/ε/	pera 'pear'	persa 'Persian'	espessa 'thick, f.'
	vela 'sail'	pelfa 'plush'	carpeta 'folder'
	ferra 'shoes'*	verda 'green, f.'	mofeta 'skunk'
	pela 'peels'*	ferma 'steady, f.'	desferra 'waste'
	<i>Dentoalveolar</i>		
	raça 'race'	tasta 'tastes/tries'*	patata 'potato'
	tassa 'cup'	rasca 'scratches'*	arrasa 'devastates'*
	tara 'defect'	tarda 'afternoon'	gatzara 'clamor'
	rata 'rat'	tasca 'task'	retrata 'portrays'*
	rasa 'flat, f.'	raspa 'scrapes'*	terrassa 'terrace'



Vowel	'CV.CV	'CVC.CV	CV'CV.CV
	<i>Labial</i>		
	passa 'step'	pasta 'pasta'	repassa 'revises'*
	para 'stops'*	parla 'speaks'*	empata 'ties'*
	pare 'father'	farta 'fed up, f.'	repara 'repairs'*
	parra 'vine'	farsa 'farce'	espasa 'sword'
	pala 'shovel'	parca 'scarce'	safata 'tray'
	<i>Dentoalveolar</i>		
	dona 'woman'	torça 'bends'*	carrossa 'carriage'
	grossa 'large, f.'	torta 'bent, f.'	carota 'mask'
	brossa 'trash'	rosca 'thread'	garota 'sea urchin'
/ɔ/	flora 'flora'	torca 'wipes'*	derrota 'defeat'
	prosa 'prose'	solta 'loose, f.'	deshora 'ungodly hour'
	<i>Labial</i>		
	pota 'leg'	porta 'door'	proposa 'proposes'
	posa 'puts'	forta 'strong'	compota 'compote'
	fora 'outside'	força 'strength'	esposa 'wife'
	vola 'flies'*	postres 'dessert'	reposa 'rests'*
	vora 'side'	porca 'sow'	enfora 'outside'
	<i>Dentoalveolar</i>		
	tota 'all, f.'	orca 'orca'	lactosa 'lactose'
/o/	rossa 'blonde, f.'	tosca 'coarse, f.'	explota 'explodes'*
	dóna 'gives'*	sorda 'deaf, f.'	estora 'rug'
	sota 'under'	sostre 'ceiling'	lectora 'reader, f.'
	torra 'toasts'*	crosta 'crust'	cotorra 'parrot'
	<i>Labial</i>		
	porra 'nightstick'	fosca 'dark, f.'	perfora 'perforates'*
	fosa 'melted, f.'	forma 'shape'	tramposa 'cheater, f.'
	fossa 'trench'	forca 'pitchfork'	espora 'spore'
	fotre 'do, pop.'	porpra 'purple'	esborra 'erases'*
	potra 'good luck'	polpa 'pulp'	rebota 'bounces'*
/u/	<i>Dentoalveolar</i>		
	sura 'floats'*	tusta 'knocks'*	astuta 'cunning, f.'
	ruta 'route'	rusca 'type of bark'	batuta 'baton'
	rusa 'Russian, f.'	turca 'Turk, f.'	lectura 'reading'
	gruta 'cave'	rústec 'rustic'	natura 'nature'
	bruta 'dirty, f.'	frustra 'frustrates'*	atura 'pauses'*
	<i>Labial</i>		
	puça 'flea'	purga 'purges'*	refusa 'refuses'*
	fura 'ferret'	fusta 'wood'	computa 'computes'*
	pura 'pure, f.'	furga 'rummages'*	amputa 'amputates'*
	fusa '32 <sup>nd</sup> note'	furta 'steals'*	embussa 'clogs up'*
	burra 'donkey, f.'	pulcre 'neat'	depura 'cleanses'*

\* = 3<sup>rd</sup> person singular, present Indicative.

f. = feminine

pop. = popular

Multiple glosses are possible for some of these words.

## Appendix B. Carrier sentences

The target words presented in Appendix A were embedded in one of these five carrier sentences (at random). The same target word was always presented in the same carrier phrase.

<i>Escriuen “X” sobre la taula.</i>	‘(They) write “X” on the table.’
<i>Sentien “X” per la finestra.</i>	‘(They) heard “X” through the window.’
<i>Dirien “X” des de l’entrada.</i>	‘(They) would say “X” from the hall.’
<i>Lletregen “X” sense parar-se.</i>	‘(They) spell “X” without pausing.’
<i>Llegeixen “X” sota la manta.</i>	‘(They) read “X” under the blanket.’

## Appendix C. Speaker judgments

Individual judgments for each of the 120 words presented that contained a stressed mid vowel. Words are divided by ExperimenterV. Numbers at the top of each category show the total number of words for which a speaker agreed with ExperimenterV, where agreement = 1, and disagreement = 0. Speaker uncertainty (“I don’t know” option) is shown as “NA”. The final row and left-most column marked “Total” show the number of total agreements per word or by speaker; thus the table’s final row indicates the speaker’s overall agreement rate as a percentage (out of 120 total words).

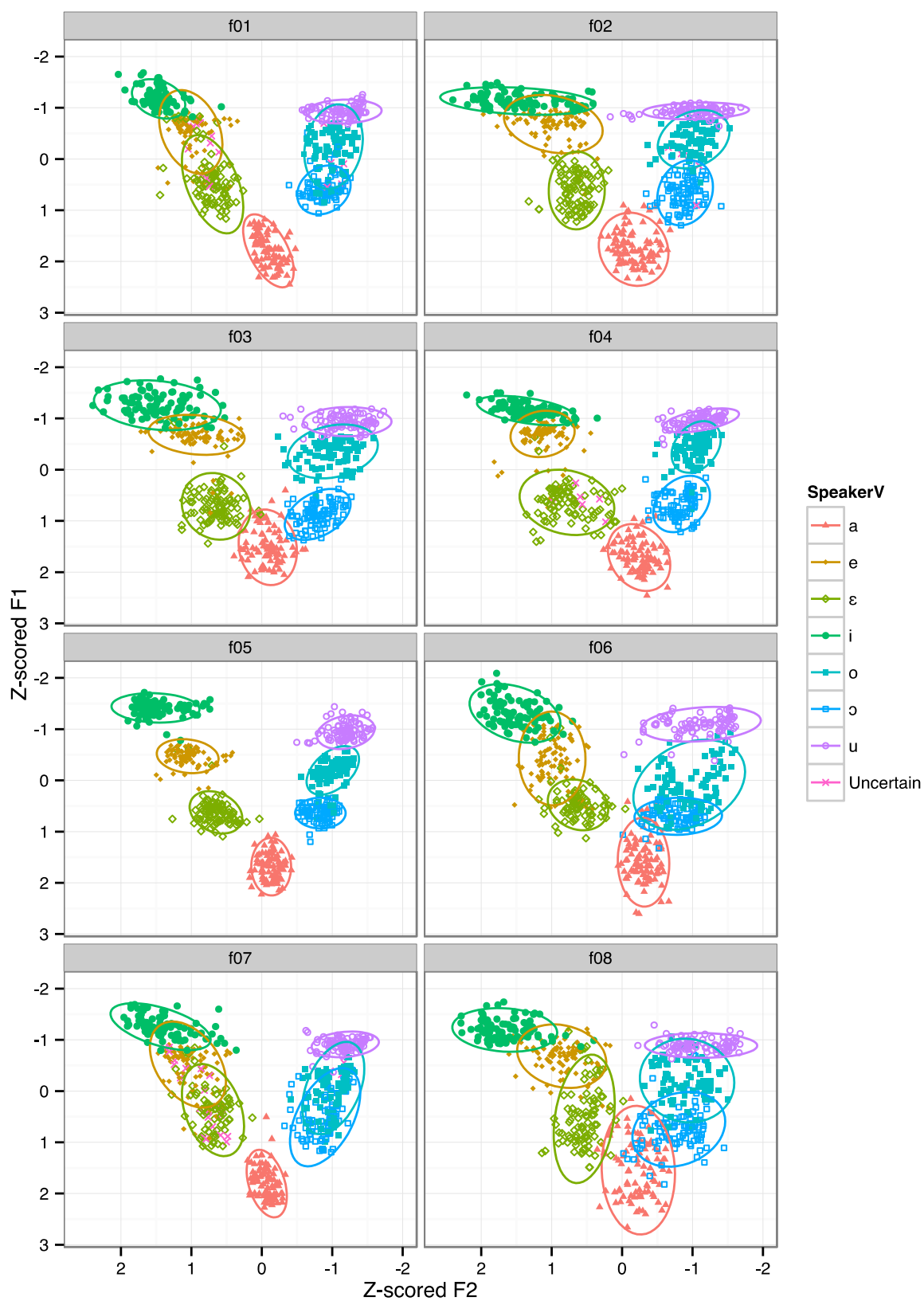
Word	f01	f02	f03	f04	f05	f06	f07	f08	f09	f10	f11	f12	f13	f14
<b>e</b>	<b>23</b>	<b>27</b>	<b>24</b>	<b>25</b>	<b>26</b>	<b>26</b>	<b>19</b>	<b>24</b>	<b>25</b>	<b>18</b>	<b>23</b>	<b>22</b>	<b>30</b>	<b>26</b>
barrera	1	1	1	1	1	1	1	1	1	1	1	1	1	1
cercle	1	1	1	1	1	1	1	1	1	1	0	1	1	1
desfeta	1	1	1	1	1	1	1	1	1	0	1	1	1	1
enceta	NA	1	0	0	1	0	0	0	1	0	0	0	1	0
esfera	1	1	1	1	1	1	1	1	1	1	1	1	1	1
espera	1	1	1	1	1	1	1	1	1	0	0	1	1	1
esvera	1	1	1	1	1	1	NA	1	1	1	1	1	1	1
fera	1	1	1	1	1	1	1	1	1	0	1	1	1	1
festa	1	1	1	1	1	1	1	1	1	0	1	1	1	1
hereta	1	1	1	1	1	1	1	1	0	1	0	0	1	1
néta	1	1	1	1	1	1	1	1	1	1	1	1	1	1
pantera	0	1	1	1	1	1	1	1	1	1	1	1	1	1
peça	1	1	1	1	1	1	1	1	1	1	1	1	1	1
pedra	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Pere	1	1	1	1	1	1	1	1	1	1	1	1	1	1
pesa	0	0	0	0	0	0	0	0	1	1	1	0	1	0
pesca	1	0	0	0	0	0	NA	0	0	1	0	0	1	0
pesta	1	1	1	1	1	1	0	1	1	1	1	1	1	1
petja	1	1	0	1	1	1	1	1	1	0	1	NA	1	1
pressa	1	1	1	1	1	1	1	1	1	1	1	1	1	1
presta	1	1	1	1	1	1	0	0	0	0	1	1	1	1
rere	1	1	1	1	1	1	NA	1	1	0	1	1	1	1
resa	0	1	0	0	0	0	0	0	0	0	0	0	1	1
resta	NA	1	1	1	1	1	1	1	1	1	1	1	1	1
sencera	NA	0	0	0	1	1	0	0	0	0	1	0	1	0
severa	0	1	1	1	1	1	NA	1	1	1	1	1	1	1

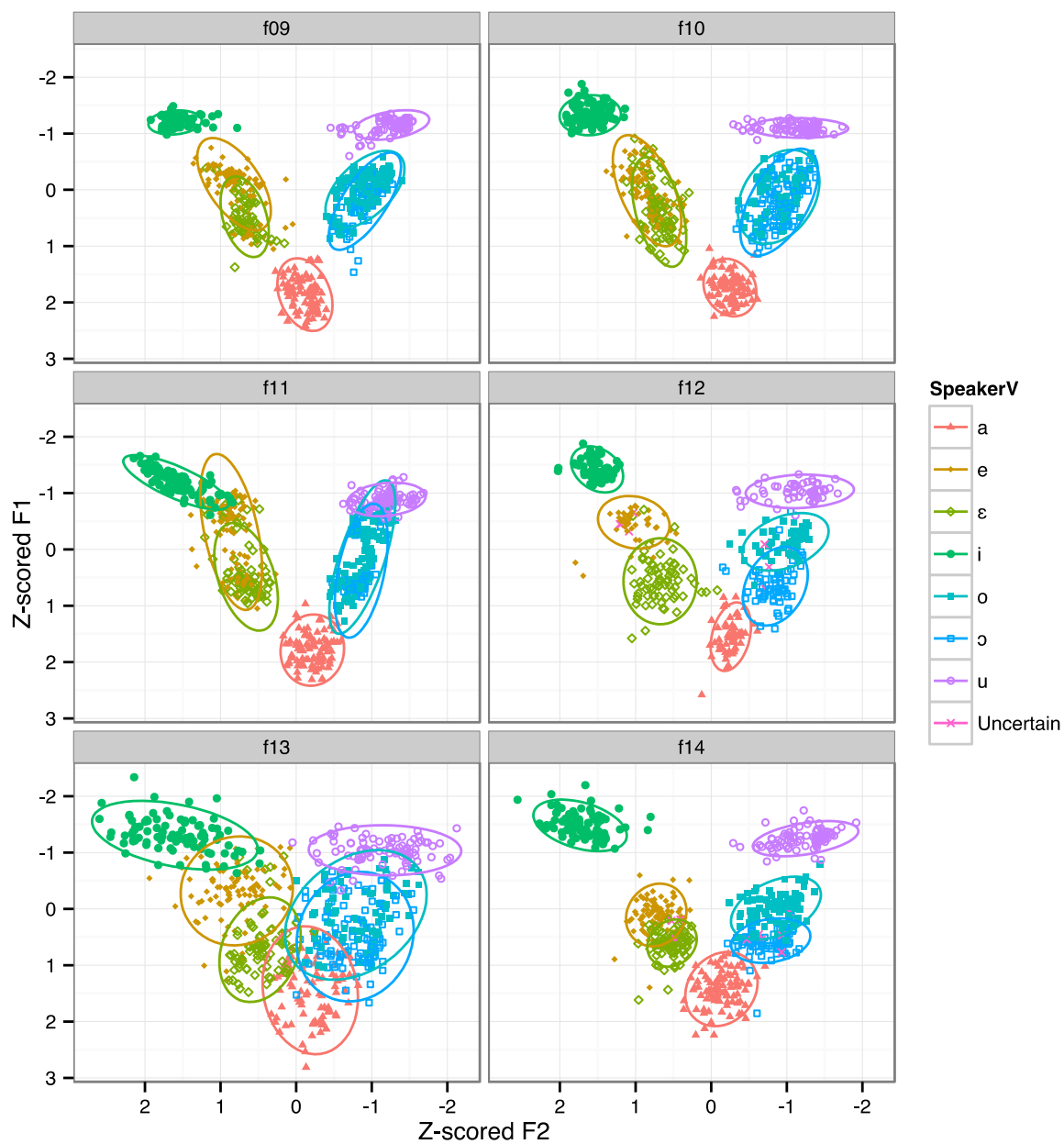
Word	f01	f02	f03	f04	f05	f06	f07	f08	f09	f10	f11	f12	f13	f14
testa	1	1	1	1	1	1	0	1	1	0	1	NA	1	1
treta	1	1	1	1	0	1	1	1	1	1	0	1	1	1
tretze	1	1	1	1	1	1	1	1	1	1	1	1	1	1
vespa	1	1	1	1	1	1	1	1	1	0	1	1	1	1
<b>o</b>	<b>26</b>	<b>27</b>	<b>27</b>	<b>27</b>	<b>28</b>	<b>28</b>	<b>23</b>	<b>26</b>	<b>22</b>	<b>14</b>	<b>23</b>	<b>24</b>	<b>14</b>	<b>28</b>
cotorra	1	1	1	1	1	1	NA	1	1	1	1	1	1	1
crosta	0	0	0	0	0	0	0	0	1	0	1	0	0	1
dóna	1	1	1	1	1	1	1	1	0	1	1	1	1	1
esborra	1	1	1	1	1	1	1	1	1	1	1	1	1	1
espora	1	1	1	1	1	1	1	1	1	1	0	1	0	1
estora	NA	1	1	1	1	1	1	1	1	0	0	0	0	1
explota	1	1	1	1	1	1	1	1	1	1	1	1	1	1
forca	1	1	1	1	1	1	1	1	0	1	1	0	0	1
forma	1	1	1	1	1	1	1	1	1	1	1	1	1	1
fosa	1	1	1	1	1	1	1	1	0	1	1	1	0	1
fosca	1	1	1	1	1	1	1	0	0	0	1	1	1	1
fossa	1	1	1	1	1	1	1	1	1	0	1	1	1	1
fotre	1	1	1	1	1	1	1	1	1	1	0	1	1	1
lactosa	1	1	1	1	1	1	1	1	1	0	1	1	1	1
lectora	1	1	1	1	1	1	1	1	1	1	0	1	1	1
orca	NA	NA	0	0	0	0	0	0	0	0	0	0	0	0
perfora	0	1	1	1	1	1	1	0	1	0	1	1	0	1
polpa	1	1	1	1	1	1	0	1	0	0	0	0	1	1
porpra	1	1	1	1	1	1	0	1	0	0	1	NA	0	1
porra	1	1	1	1	1	1	0	1	1	0	1	1	0	1
potra	1	1	1	1	1	1	1	1	1	0	1	1	0	NA
rebota	1	1	1	1	1	1	1	1	1	0	1	1	0	1
rossa	1	1	1	1	1	1	1	1	1	0	1	1	0	1
sorda	1	1	1	1	1	1	1	1	1	0	1	1	0	1
sostre	1	1	1	1	1	1	1	1	1	1	1	1	0	1
sota	1	1	1	1	1	1	1	1	1	1	1	1	1	1
torra	1	1	1	1	1	1	1	1	1	0	1	1	0	1
tosca	1	0	0	0	1	1	0	1	1	1	1	1	0	1
tota	1	1	1	1	1	1	1	1	0	0	0	1	1	1
tramposa	1	1	1	1	1	1	1	1	1	1	1	1	1	1
<b>ε</b>	<b>27</b>	<b>27</b>	<b>25</b>	<b>25</b>	<b>28</b>	<b>26</b>	<b>21</b>	<b>25</b>	<b>16</b>	<b>18</b>	<b>17</b>	<b>30</b>	<b>26</b>	<b>25</b>
aterra	1	1	1	1	1	1	1	1	0	1	0	1	1	1
cabeça	1	1	1	1	1	1	1	1	1	1	0	1	0	1
carpeta	0	1	1	1	1	1	1	1	1	0	0	1	1	1
cera	1	1	1	1	1	1	1	1	1	1	1	1	1	1
certa	1	1	1	1	1	1	1	1	0	1	0	1	1	1
Cesca	1	1	1	1	1	1	1	1	1	0	0	1	1	1
cresta	1	0	0	0	0	1	0	0	0	0	0	1	1	0
desferra	1	1	1	NA	1	0	1	0	0	0	1	1	1	1
empresa	1	1	1	1	1	1	0	1	1	0	0	1	1	1
espessa	1	1	0	1	1	1	1	0	1	0	0	1	1	1
estesa	1	1	1	1	1	1	1	1	0	0	0	1	1	0
estreta	1	1	0	1	1	1	1	1	0	0	1	1	1	1
ferma	1	1	1	1	1	1	1	1	0	1	1	1	1	1
ferra	1	1	1	NA	1	1	0	1	0	1	1	1	1	NA
fresca	1	1	1	1	1	1	1	1	1	1	1	1	1	1
fressa	1	0	0	0	1	0	NA	0	0	1	0	1	0	0

Word	f01	f02	f03	f04	f05	f06	f07	f08	f09	f10	f11	f12	f13	f14
mestressa	1	1	1	1	1	1	1	1	1	1	1	1	1	1
mofeta	1	1	1	1	1	1	1	1	1	1	0	1	1	1
neta	1	1	1	1	1	1	0	1	1	1	1	1	0	1
pela	1	1	1	1	1	1	NA	1	1	1	1	1	1	1
pelfa	1	1	1	1	1	1	1	1	1	1	1	1	1	1
pera	NA	1	1	1	1	1	1	1	1	0	1	1	1	1
perdre	1	1	1	1	1	1	1	1	0	0	1	1	1	1
persa	1	1	1	1	1	1	NA	1	1	1	1	1	1	1
peta	1	1	1	1	1	1	0	1	0	0	1	1	1	1
setze	0	0	0	0	0	0	0	0	0	0	0	1	0	0
tela	1	1	1	1	1	1	1	1	0	1	1	1	1	1
vela	1	1	1	1	1	0	1	1	1	1	1	1	1	1
verda	1	1	1	1	1	1	1	1	0	1	1	1	1	1
zeta	1	1	1	1	1	1	1	1	1	1	0	1	1	1
<b>ɔ</b>	<b>26</b>	<b>29</b>	<b>29</b>	<b>28</b>	<b>26</b>	<b>28</b>	<b>22</b>	<b>30</b>	<b>13</b>	<b>18</b>	<b>11</b>	<b>28</b>	<b>19</b>	<b>28</b>
brossa	1	1	1	1	1	1	1	1	1	1	0	1	1	1
carota	1	1	1	1	1	1	1	1	1	1	0	1	0	1
carrossa	1	1	1	1	1	1	1	1	1	0	0	1	1	1
compota	0	1	1	1	1	1	1	1	0	0	0	1	1	1
derrota	1	1	1	1	1	1	1	1	0	1	0	1	1	1
deshora	1	1	1	1	1	1	1	1	1	0	1	1	1	1
dona	1	1	1	1	1	1	0	1	0	1	0	1	1	1
enfora	1	1	1	1	1	1	1	1	1	1	1	1	0	1
esposa	1	1	1	1	1	1	0	1	0	1	0	1	0	1
flora	1	NA	0	1	0	0	1	1	0	1	1	0	1	0
fora	1	1	1	1	1	1	1	1	1	1	1	1	0	1
força	1	1	1	1	1	1	1	1	1	1	0	1	1	1
forta	1	1	1	1	1	1	1	1	1	0	0	1	1	1
garota	1	1	1	1	1	1	1	1	0	1	0	1	0	NA
grossa	1	1	1	1	1	1	1	1	0	0	1	1	1	1
porca	1	1	1	1	1	1	0	1	1	0	0	1	1	1
porta	1	1	1	1	1	1	1	1	0	0	1	1	1	1
posa	1	1	1	1	1	1	0	1	1	0	0	1	1	1
postres	NA	1	1	1	1	1	1	1	0	0	0	1	1	1
pota	1	1	1	1	1	1	1	1	1	0	1	1	1	1
proposa	0	1	1	1	1	1	1	1	0	0	0	1	1	1
prosa	1	1	1	1	1	1	0	1	0	0	0	1	0	1
reposa	1	1	1	0	0	1	0	1	0	1	0	1	0	1
rosca	1	1	1	1	1	1	1	1	1	1	0	1	0	1
solta	1	1	1	1	1	1	1	1	1	1	1	1	0	1
torca	1	1	1	1	0	1	0	1	0	1	1	NA	0	1
torça	1	1	1	1	1	1	0	1	0	1	0	1	0	1
torta	1	1	1	1	1	1	1	1	0	1	0	1	1	1
vola	1	1	1	0	0	0	1	1	0	1	1	1	1	1
vora	0	1	1	1	1	1	1	1	0	1	1	1	1	1
<b>Total</b>	<b>102</b>	<b>110</b>	<b>105</b>	<b>105</b>	<b>108</b>	<b>108</b>	<b>85</b>	<b>105</b>	<b>76</b>	<b>68</b>	<b>74</b>	<b>104</b>	<b>89</b>	<b>107</b>

**Appendix D. By-speaker Vowel Plots**

Individual vowel plots. Vowels are plotted taking SpeakerV into account, using normalized (z-scored) F1, F2 values for maximum visual comparability across speakers. Ellipses represent 95% confidence interval based on SpeakerV, and exclude any Uncertain tokens. Each point represents one repetition of a target vowel.





## Appendix E. Results of mixed-effects modeling

The default level of categorical variables is shown in [brackets].

### Model of normalized F1 values

	<b><math>\beta</math> Estimate</b>	<b>Std. Error</b>	<b><i>t</i></b>	<b><i>p</i></b>
Intercept	0.3325	0.2112	1.5740	0.1181
Pair[front]	-0.2728	0.1082	-2.5200	0.0131
Height[low]	0.1653	0.0219	7.5590	0.0000
Context[labial]	-0.1524	0.1082	-1.4080	0.1618
Repetition	-0.0200	0.0154	-1.2960	0.2170
Syllable count	-0.0322	0.0931	-0.3450	0.7304
Syllable type	0.0133	0.0931	0.1430	0.8868
Pair[front] $\times$ Height[low]	0.0603	0.0261	2.3060	0.0211
Pair[front] $\times$ Context[labial]	0.1869	0.1520	1.2290	0.2215
Height[low] $\times$ Context [labial]	0.0034	0.0258	0.1330	0.8940

### Model of normalized F2 values

	<b><math>\beta</math> Estimate</b>	<b>Std. Error</b>	<b><i>t</i></b>	<b><i>p</i></b>
Intercept	-0.6984	0.0721	-9.6880	~ 0
Pair = front	1.5270	0.0372	41.1120	~ 0
Height = low	0.0122	0.0132	0.9230	0.3560
Context = labial	-0.3649	0.0371	-9.8270	~ 0
Repetition	-0.0026	0.0076	-0.3380	0.7410
Syllable count	-0.0130	0.0315	-0.4140	0.6800
Syllable type	0.0384	0.0315	1.2220	0.2240
Pair[front] $\times$ Height[low]	-0.0907	0.0158	-5.7420	0.0000
Pair[front] $\times$ Context[labial]	0.3613	0.0514	7.0340	0.0000
Height[low] $\times$ Context[labial]	0.0104	0.0156	0.6630	0.5070



## References

- Adank, P., Smits, R., & Van Hout, R. (2004). A comparison of vowel normalization procedures for language variation research. *The Journal of the Acoustical Society of America*, 116(5), 3099–3107.
- Alcover, A. M., & Moll, F. de B. (1978). *Diccionari Català Valencià Balear*. Palma de Mallorca.
- Alkire, T., & Rosen, C. (2010). *Romance Languages: A Historical Introduction*. Cambridge: Cambridge University Press.
- Amengual, M. (2014). The perception and production of language-specific mid-vowel contrasts: Shifting the focus to the bilingual individual in early language input conditions. *International Journal of Bilingualism*.
- Amengual, M., & Chamorro, P. (2015). The Effects of Language Dominance in the Perception and Production of the Galician Mid Vowel Contrasts. *Phonetica*, 207–236.  
<http://doi.org/10.1159/000439406>
- Avolio, F. (1995). *Bommèspre: profilo linguistico dell'Italia centro-meridionale*. San Severo, Italy: Gerni.
- Badia i Margarit, A. M. (1968). Algunes normes de prosòdia catalana, segons les rimes de Carles Riba. In A. M. Badia i Margarit, *Sons i fonemes de la llengua catalana*. Barcelona: Publicacions de la Universitat de Barcelona.
- Badia i Margarit, A. M. (1969a). Algunes mostres de les igualacions E = e i O = o en el català parlat de Barcelona. In A. M. Badia i Margarit, *Sons i fonemes de la llengua catalana* (pp. 97–103). Barcelona: Publicacions de la Universitat de Barcelona.

- Badia i Margarit, A. M. (1969b). Rimas vocálicas anómalas en el Cancionero Popular Catalán. In A. M. Badia i Margarit, *Sons i fonemes de la llengua catalana* (pp. 191–208). Barcelona: Publicacions de la Universitat de Barcelona.
- Badia i Margarit, A. M. (1970a). Les oppositions phonologiques E/e et O/o du catalan dans les rimes des poètes modernes. In A. M. Badia i Margarit, *Sons i fonemes de la llengua catalana* (pp. 105–140). Barcelona: Publicacions de la Universitat de Barcelona.
- Badia i Margarit, A. M. (1970b). Les vocals tòniques e i o en el català de Barcelona [1970]. Assaig d'anàlisi fonològica de la situació actual. In A. M. Badia i Margarit, *Sons i fonemes de la llengua catalana* (pp. 141–190). Barcelona: Publicacions de la Universitat de Barcelona.
- Badia i Margarit, A. M. (1981). *Gramàtica històrica catalana*. València: Tres i quatre.
- Badia Margarit, A. M. (1966). Predominio de las vocales abiertas e y o en el catalán de Barcelona. *Revista de Filología Española*, 49(1/4), 315–320.
- Bates, D., & Maechler, M. (2009). lme4: Linear mixed-effects models using Eigen and R syntax (Version 0.999375-32). R package version 0.999375-32).
- Birdsong, D., Gertken, L. M., & Amengual, M. (2012). Bilingual language profile: An easy-to-use instrument to assess bilingualism. *COERLL, University of Texas at Austin*.
- Bloch, B., & Trager, G. L. (1942). *Outline of linguistic analysis*. Linguistic Society of America at the Waverly Press, Inc.
- Bloomfield, L. (1933). *Language*. University of Chicago Press.
- Boersma, P., & Weenink, D. (2015). Praat: Doing phonetics by computer [Computer program], Version 5.4.08 (Version 5.4.08). Retrieved from <http://www.fon.hum.uva.nl/praat/>

- Bosch, L., Costa, A., & Sebastián-Gallés, N. (2000). First and second language vowel perception in early bilinguals. *European Journal of Cognitive Psychology*, 12(2), 189–221.
- Bosch, L., & Ramon-Casas, M. (2011). Variability in vowel production by bilingual speakers: Can input properties hinder the early stabilization of contrastive categories? *Journal of Phonetics*, 39(4), 514–526.
- Bosch, L., & Sebastián-Gallés, N. (2003). Simultaneous bilingualism and the perception of a language-specific vowel contrast in the first year of life. *Language and Speech*, 46(2–3), 217–243.
- Bruguera i Talleda, J. (2004). *Diccionari ortogràfic i de pronuncia*. Barcelona: Enciclopèdia Catalana.
- Bybee, J. L. (2010). *Language, usage and cognition*. Cambridge; New York: Cambridge University Press.
- Cabré, T. (2009). Vowel reduction and vowel harmony in Eastern Catalan loanword phonology. In M. C. Vigário, S. Frota, & M. J. Freitas (Eds.), *Phonetics and phonology: interactions and interrelations* (pp. 267–285). Amsterdam; Philadelphia: John Benjamins Publishing Company.
- Chen, Y. (2008). The acoustic realization of vowels of Shanghai Chinese. *Journal of Phonetics*, 36(4), 629–648.
- Coromines, J. (1974). *Lleures i converses d'un filòleg*. Barcelona: Club Editor.
- De Boer, B. (2011). First formant difference for /i/ and /u/: A cross-linguistic study and an explanation. *Journal of Phonetics*, 39(1), 110–114.

- Diehl, R. L. (2008). Acoustic and auditory phonetics: the adaptive design of speech sound systems. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 363(1493), 965–978.
- Dobson, E. J. (1968). *English Pronunciation 1500-1700* (2nd ed). Oxford: Clarendon Press.
- Dufour, S., Nguyen, N., & Frauenfelder, U. H. (2007). The perception of phonemic contrasts in a non-native dialect. *The Journal of the Acoustical Society of America*, 121(4), EL131–136.
- Escudero, P., Boersma, P., Rauber, A. S., & Bion, R. A. H. (2009). A cross-dialect acoustic description of vowels: Brazilian and European Portuguese. *Journal of the Acoustical Society of America*, 126(3), 1379–1393.
- Fabra, P. (1906). Les e toniques du catalan. *Revue Hispanique*, 15(47), 9–23.
- Fagyal, Z., Hassa, S., & Ngom, F. (2002). L’opposition [e]-[ɛ] en syllabes ouvertes de fin de mot en français parisien: étude acoustique préliminaire. In *Actes des Journées d’Etudes sur la Parole* (pp. 165–168). Nancy, France.
- Fagyal, Z., Kibbee, D., & Jenkins, F. (2006). *French: A linguistic introduction*. Cambridge University Press.
- Fleiss, J. L. (1971). Measuring nominal scale agreement among many raters. *Psychological Bulletin*, 76(5), 378–382.
- Flemming, E. (1995). *Auditory Representations in Phonology* (Ph.D. Dissertation). UCLA, Los Angeles, California.
- Fougeron, C., & Audibert, N. (2011). Testing various metrics for the description of vowel distortion in dysarthria. In *Proceedings of the 17th International Congress of Phonetic Sciences* (pp. 1–4). Hong Kong.

- Gamer, M., Lemon, J., Fellows, I., & Singh, P. (2012). Irr: Various Coefficients of Interrater Reliability and Agreement. *R Package Version 0.84*.
- Goldsmith, J. A. (1995). Phonological Theory. In J. A. Goldsmith (Ed.), *The Handbook of Phonological Theory* (pp. 1–23). Cambridge, MA: Blackwell Publishers.
- Hall, K. C. (2009). *A Probabilistic Model of Phonological Relationships from Contrast to Allophony* (Ph.D. Dissertation). The Ohio State University, Columbus, OH.
- Hall, K. C. (2013). A typology of intermediate phonological relationships. *The Linguistic Review*, 30(2), 215–275.
- Hall, K. C., & Hume, E. (2015, August). *Modeling Perceived Similarity: The influence of phonetics, phonology and frequency on the perception of French vowels*. Manuscript, submitted August 2015.
- Hall-Lew, L. (2010). Improved representation of variance in measures of vowel merger. In *Proceedings of Meetings on Acoustics* (Vol. 9, p. 60002). Acoustical Society of America. Retrieved from <http://scitation.aip.org/content/asa/journal/poma/9/1/10.1121/1.3460625>
- Hartigan, J. A., & Wong, M. A. (1979). A K-means clustering algorithm. *Applied Statistics*, 28, 100–108.
- Hay, J., Warren, P., & Drager, K. (2006). Factors influencing speech perception in the context of a merger-in-progress. *Journal of Phonetics*, 34(4), 458–484.
- Hockett, C. F. (1955). *A manual of phonology* (Vol. 4). Baltimore: Waverly Press.
- Hockett, C. F. (1966). *The quantification of functional load: A linguistic problem* (Memorandum No. RM-5168-PR).
- Hualde, J. I. (2004). Quasi-phonemic contrasts in Spanish. In *WCCFL 23: Proceedings of the 23rd West Coast Conference on Formal Linguistics* (pp. 374–398).

- Hyman, L. M. (1975). *Phonology: Theory and Analysis*. New York: Holt, Rinehart and Winston.
- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2013). lmerTest: Tests for random and fixed effects for linear mixed effect models (lmer objects of lme4 package). *R Package Version*, 2–0.
- Ladd, D. R. (2006). “Distinctive phones” in surface representation. In L. Goldstein, D. H. Whalen, & C. T. Best (Eds.), *Laboratory Phonology 8* (pp. 3–26). Berlin; New York: Mouton de Gruyter.
- Landick, M. (1995). The mid-vowels in figures: hard facts. *French Review*, 88–102.
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33(1), 159–174.
- Lindblom, B. (1986). Phonetic universals in vowel systems. In J. J. Ohala & J. J. Jaeger (Eds.), *Experimental phonology* (pp. 13–44). Orlando, Florida: Academic Press.
- Lobanov, B. M. (1971). Classification of Russian vowels spoken by different speakers. *The Journal of the Acoustical Society of America*, 49(2B), 606–608.
- Luick, K. (1940). *Historische Grammatik der englischen Sprache*. Leipzig: C. H. Tauchnitz.
- Mascaró, J. (1984). Aramon-aramònic: una regla fonològica catalana no descoberta fins ara. In *Estudis de llengua i literatura catalanes oferts a R. Aramon i Serra* (pp. 123–129). Barcelona: Curial.
- Mascaró, J. (2002). Reducció vocàlica. In M.-R. L. Solà, J. Mascaró, & M. Pérez Saldanya (Eds.), *Gramàtica Del Català Contemporani* (pp. 89–123). Barcelona, Spain: Empúries.
- Mascaró, J. (2008). La distribució de les vocals mitjanes tòniques en català central. In *Caplletra* (pp. 79–102).

- Mascaró, J. (2010). Efectes accentuals i sil·làbics sobre la distribució de les vocals mitjanes en català central. In M.-R. Lloret & C. Pons (Eds.), *Noves aproximacions a la fonologia i la morfologia del català* (pp. 405–426). Alacant: Institut Interuniversitari de Filologia Valenciana.
- Moll, F. de B. (1991). *Gramàtica històrica catalana*. Universitat de València.
- Mora, J. C., Keidel, J. L., & Flege, J. E. (2011). Why are the Catalan contrasts between /e/-/ɛ/ and /o/-/ɔ/ so difficult for even early Spanish-Catalan bilinguals to perceive? In M. Wrembel, M. Kul, & K. Dziubalska-Kołaczyk (Eds.), *Achievements and perspectives in the acquisition of second language speech: New Sounds 2010. Volume II* (pp. 183–193). Frankfurt am Mein: Peter Lang.
- Mora, J. C., Keidel, J. L., & Flege, J. E. (2015). Effects of Spanish use on the production of Catalan vowels by early Spanish-Catalan bilinguals. In J. Romero & M. Riera (Eds.), *The Phonetics-Phonology Interface: Sounds, Representations, Methodologies* (pp. 33–53). Amsterdam: John Benjamins.
- Mora, J. C., & Nadeu, M. (2012). L2 effects on the perception and production of a native vowel contrast in early bilinguals. *International Journal of Bilingualism*, 16(4), 484–500.
- Nadeu, M. (2014). Stress- and speech rate-induced vowel quality variation in Catalan and Spanish. *Journal of Phonetics*, 46, 1–22. <http://doi.org/10.1016/j.wocn.2014.05.003>
- Nadeu, M. (2016). Phonetic and phonological vowel reduction in Central Catalan. *Journal of the International Phonetic Association*, 46(1), 33–60. <http://doi.org/10.1017/S002510031500016X>

- Navarra, J., Sebastián-Gallés, N., & Soto-Faraco, S. (2005). The perception of second language sounds in early bilinguals: new evidence from an implicit measure. *Journal of Experimental Psychology: Human Perception and Performance*, 31(5), 912–918.
- Pallier, C., Bosch, L., & Sebastián-Gallés, N. (1997). A limit on behavioral plasticity in speech perception. *Cognition*, 64(3), B9–B17.
- Pallier, C., Colomé, A., & Sebastián-Gallés, N. (2001). The influence of native-language phonology on lexical access: Exemplar-based versus abstract lexical entries. *Psychological Science*, 12(6), 445–449.
- Pi-Mallarach, J. (2001). Els neologismes com a indicadors de l'evolució fonològica de la llengua. El cas de les vocals mitjanes en el català central actual. In J. Brumme (Ed.), *La historia de los lenguajes iberorrománicos de especialidad: la divulgación de la ciencia* (pp. 317–331). Barcelona: Institut Universitari de Lingüística Aplicada.
- Pi-Mallarach, J. (2006). La pronúncia oberta o tancada de les oo en els neologismes nominals del català central: ¿És la presentació d'una tendència innovadora o bé la confirmació d'una manifestació antiga? In S. Martí, M. Cabré, F. Feliu, N. Iglesias, & D. Prats (Eds.), *Actes del Tretzè Col·loqui Internacional de Llengua i Literatura Catalanes* (Vol. II, pp. 311–325). Barcelona: Publicacions de l'Abadia de Montserrat.
- Pi-Mallarach, J. (2007). L'actualitat de Fabra en l'adaptació fonològica de les vocals mitjanes anteriors en mots de nova creació. In A. Montserrat & O. Cubells (Eds.), *Entorn i vigència de l'obra de Fabra: Actes del 2n Col·loqui Internacional La Lingüística de Pompeu Fabra* (pp. 229–238). Valls: Cossetània.
- R Core Team. (2000). *R Language Definition*. Available from CRAN sites. Retrieved from <ftp://155.232.191.133/cran/doc/manuals/r-devel/R-lang.pdf>



- Rafel i Fontanals, J. (1980). Dades sobre la freqüència de les unitats fonològiques del català. *Estudis Universitaris Catalans*, (24), 473–496.
- Ramon-Casas, M., Swingle, D., Sebastián-Gallés, N., & Bosch, L. (2009). Vowel categorization during word recognition in bilingual toddlers. *Cognitive Psychology*, 59(1), 96–121.
- Rasico, P. D. (1981). *Preliterary Catalan Historical Phonology* (Ph.D. Dissertation). Indiana University, Department of Spanish and Portuguese., Bloomington, Indiana.
- Recasens, D. (1985). Coarticulatory patterns and degrees of coarticulatory resistance in Catalan CV sequences. *Language and Speech*, 28(2), 97–114.
- Recasens, D. (1993). *Fonètica i fonologia*. Barcelona, Spain: Enciclopèdia Catalana.
- Recasens, D. (2014). *Fonètica i fonologia experimentals del català. Vocals i consonants*. Barcelona: Institut d'Estudis Catalans.
- Recasens, D., & Espinosa, A. (2009). Dispersion and variability in Catalan five and six peripheral vowel systems. *Speech Communication*, 51(3), 240–258.
- Renwick, M. E. L. (2011). Phoneme Type Frequency in Romanian. *University of Pennsylvania Working Papers in Linguistics*, 17(1), 1–22.
- Renwick, M. E. L. (2014). *The Phonetics and Phonology of Contrast: The Case of the Romanian Vowel System*. Berlin, Boston: De Gruyter Mouton.
- Renwick, M. E. L., & Ladd, D. R. (Under review). *Phonetic distinctiveness vs. lexical contrastiveness in non-robust phonemic contrasts*. Under review at *Laboratory Phonology*.
- Schwartz, J.-L., Boë, L.-J., Vallée, N., & Abry, C. (1997). The dispersion-focalization theory of vowel systems. *Journal of Phonetics*, 25(3), 255–286.

- Scobbie, J. M., & Stuart-Smith, J. (2008). Quasi-phonemic contrast and the fuzzy inventory: Examples from Scottish English. In P. Avery, B. E. Dresher, & K. Rice (Eds.), *Contrast in Phonology: Theory, Perception, Acquisition* (pp. 87–114). Berlin: de Gruyter.
- Sebastián-Gallés, N., Echeverría, S., & Bosch, L. (2005). The influence of initial exposure on lexical representation: Comparing early and simultaneous bilinguals. *Journal of Memory and Language*, 52(2), 240–255.
- Sebastian-Gallés, N., Rodríguez-Fornells, A., de Diego-Balaguer, R., & Díaz, B. (2006). First- and second-language phonological representations in the mental lexicon. *Journal of Cognitive Neuroscience*, 18(8), 1277–1291.
- Sebastián-Gallés, N., & Soto-Faraco, S. (1999). Online processing of native and non-native phonemic contrasts in early bilinguals. *Cognition*, 72(2), 111–123.
- Sebastián-Gallés, N., Vera-Constán, F., Larsson, J. P., Costa, A., & Deco, G. (2009). Lexical plasticity in early bilinguals does not alter phoneme categories: II. Experimental evidence. *Journal of Cognitive Neuroscience*, 21(12), 2343–2357.
- Simonet, M. (2011). Production of a Catalan-specific vowel contrast by early Spanish-Catalan bilinguals. *Phonetica*, 68(1–2), 88–110.
- Simonet, M. (2014). Phonetic consequences of dynamic cross-linguistic interference in proficient bilinguals. *Journal of Phonetics*, 43, 26–37.
- Sloos, M. (2013). The reversal of the BÄREN-BEEREN merger in Austrian Standard German. *The Mental Lexicon*, 8(3), 353–371.
- Stevens, K. N., & Keyser, S. J. (2010). Quantal theory, enhancement and overlap. *Journal of Phonetics*, 38(1), 10–19.

- Surendran, D., & Niyogi, P. (2006). Quantifying the functional load of phonemic oppositions, distinctive features, and suprasegmentals. In O. Nedergaard Thomsen (Ed.), *Competing models of linguistic change: evolution and beyond* (pp. 43–58). Amsterdam; Philadelphia: John Benjamins Publishing Company.
- Tabain, M. (2009). An EPG study of the alveolar vs. retroflex apical contrast in Central Arrernte. *Journal of Phonetics*, 37(4), 486–501.
- Thurber, B. A. (2011). Voicing of Initial Interdental Fricatives in Early Middle English Function Words. *Journal of Germanic Linguistics*, 23(1), 65–81.
- Trubetzkoy, N. S. (1939). *Principles of Phonology*. (C. Baltaxe, Trans.). Berkeley: University of California Press.
- Vincent, N. (1988). Italian. In M. Harris & N. Vincent (Eds.), *The Romance Languages* (pp. 279–313). New York: Oxford University Press.
- Wedel, A., Jackson, S., & Kaplan, A. (2013). Functional load and the lexicon: Evidence that syntactic category and frequency relationships in minimal lemma pairs predict the loss of phoneme contrasts in language change. *Language and Speech*, 56(3), 395–417.
- Wetzels, W. L. (1992). Mid vowel neutralization in Brazilian Portuguese. *Cadernos de Estudos Lingüísticos*, 23.
- Wheeler, M. (2005). *The phonology of Catalan*. Oxford: Oxford University Press.