Corresponding author:
Margaret E. L. Renwick, University of Georgia, Department of Linguistics, 240 Gilbert Hall, Athens GA 30602, USA
Email: mrenwick@uga.edu

A survey of phonological mid-vowel intuitions in Central Catalan

Margaret E. L. Renwick

University of Georgia, Athens, GA

Marianna Nadeu

Independent researcher

Abstract

Catalan, like other Romance languages, has two pairs of phonemic mid vowels (/be/ 'well' vs. /bɛ/ 'lamb'; /os/ 'bear' vs. /ɔs/ 'bone'). However, these contrasts do not function like others in the language: they are partially phonologically conditioned, and evidence shows that words may be pronounced with different mid vowels by speakers of the same variety or even by the same speaker. Spanish may influence this instability, as L1 Spanish Catalan-Spanish bilinguals struggle to perceive and produce the contrast. This paper investigates the mid vowel contrasts in an Internet survey of vowel height judgments in 220 words by 146 Central Catalan-speaking individuals who also self-reported their language history. Results confirm that certain phonological contexts condition mid vowel height, typically favoring low mid judgments; where phonological conditioning occurs, speakers judge quality with increased consistency and confidence. Many words lacking phonological conditioning environments, however, are variable across speakers. Bilingualism levels and age have an effect: among Catalan-dominant participants, choice of mid vowel is affected by age, while participants with the highest Catalan dominance have greatest confidence in their intuitions. Variably-judged words are also phonetically variable, indicating a word-specific association between strength of phonological representation and realization.

Keywords: phonological contrast; phonological conditioning; Catalan; language dominance; mid vowels

1. Introduction

This paper presents a study examining phonological intuitions of native Central Catalan speakers. Specifically, we focus on native speaker judgments of vowel height in Catalan mid vowels /e ɛ ɔ o/ in a particular set of lexical items. Our study is neither a production task nor a canonical perception task, but rather it directly queries judgments of vowel height among literate speakers of Catalan. This technique permits the collection of data from a large number of speakers over a short period of time, using an internet-based platform.

Many psycholinguistic studies have focused on Catalan-Spanish bilingual speakers' ability to perceive the mid vowel contrasts of Catalan (Bosch, Costa, & Sebastián-Gallés, 2000; Bosch & Ramon-Casas, 2011; Navarra, Sebastián-Gallés, & Soto-Faraco, 2005; Pallier, Bosch, & Sebastián-Gallés, 1997; Pallier, Colomé, & Sebastián-Gallés, 2001; Sebastián-Gallés, Vera-Constán, Larsson, Costa, & Deco, 2009), often as a function of their age or order of acquisition (Bosch & Sebastián-Gallés, 2003; Ramon-Casas, Swingley, Sebastián-Gallés, & Bosch, 2009; Sebastián-Gallés & Soto-Faraco, 1999), arguing that initial exposure to Spanish reduces sensitivity to them, because that language does not share the mid vowel contrasts. In spite of this, recent research shows that monolingual Spanish speakers are sensitive to the phonetic distinction between low and high mid vowels (due to phonologically conditioned vowel laxing) associated to a singular-plural distinction in some varieties of Spanish (Henriksen, 2017).[1]

Our line of work, on the other hand, considers the possibility that beyond the effects of bilingualism on vowel acquisition and perception, the lexical specification of mid vowels in many Catalan words is inherently variable, even among Catalan-dominant speakers. At least as early as 1906, variation in the realization of mid vowels in Catalan has been attested especially in learned words, but also in other words (Fabra, 1906). At the same time, in many words, mid vowel quality is claimed to be predictable based on surrounding phonological characteristics (Wheeler, 2005). Together, we argue, these two properties may indicate a relatively weaker phonemic distinction between the high and low mid vowels (also *close* and *open*, respectively) in Catalan with respect to other vowels, creating a scenario in which speaker awareness of the contrast may be reduced, or variable, because it only rarely signals a lexical distinction. Such a scenario is supported by recent results showing that the mid vowel distinctions are not perceived as categorically as others in Catalan, though this is affected by language use (Mora, Keidel, & Flege, 2011). Furthermore, the instability of mid vowel contrasts in other Romance languages, including French (Fagyal, Kibbee, & Jenkins, 2006; Landick, 1995; Trubetzkoy, 1939) and Italian (Renwick & Ladd, 2016), suggests that the influence of Spanish may not be the only factor contributing to variability in mid vowel production in Catalan.

The present paper extends investigations of mid vowel variation in Catalan from the speaker's perspective, incorporating phonological factors into a survey of vowel height intuitions that also queries participants' confidence in judgment and their own abilities to produce and perceive the mid vowel contrasts. Our results confirm that many words receive consistent judgments, as predicted by their morphophonological properties, but in

a subset of words, judgments are highly variable. We also find that respondents' confidence, adherence to prescriptive models (based on etymology and on phonological conditioning), and self-reported use of the contrast is related to their level of dominance in Catalan vs. Spanish.

1.1 Survey-based studies of phonological intuition

In previous work (Nadeu & Renwick, 2016), we have shown that individual Catalan-dominant speakers may misjudge the mid vowel that they use in specific words, and/or alternate between use of the close and open mid vowels across repetitions of a word. In this complementary study, we investigate whether certain words are more variable in mid vowel height across individual speakers. Although speakers in the first study did sometimes misjudge vowel height in their own speech, they were largely accurate; thus we suppose that participants' judgments in the present study are similarly accurate, though they also reported certainty levels for each judgment.

Linguists often make judgments about language, based on their own intuition, or by listening to others' pronunciation, sentence structure, lexical choices, etc. However, quantitative studies relying on participants' explicit judgments of phonological intuition are relatively rare. Below, we situate this project among studies that rely on linguistic judgments, and describe other work that has collected judgments similarly to the present study.

Anecdotal observations by theoreticians, alongside native speaker intuitions, have long been an anchor of linguistic argumentation about the nature of sound patterns, syntax (Schütze, 1996), and other aspects of grammar. However, dependence on the sole intuition of a researcher, or dictionary writer, may obscure variation across native speakers, an issue that can be mitigated by techniques in which stimuli are presented to a sufficiently large population that is naïve to the purpose of the study, in a random order, producing results that support tests of statistical significance (Wasow & Arnold, 2005). Speaker preferences, given explicitly or implicitly, reveal important native-speaker knowledge, for instance of morphophonological generalizations (Berko, 1958), or phonological well-formedness based on phonotactics (Hay, Pierrehumbert, & Beckman, 2003), wordlikeness (S. A. Frisch, Large, & Pisoni, 2000; Stefan A. Frisch & Zawaydeh, 2001), and syllable structure (Côté & Kharlamov, 2010; Tilsen & Cohn, 2016). The fact that intuitions reflect actual linguistic performance, particularly in the case of well-attested and robust contrasts, supports the validity of using judgments, even in the absence of corresponding production data, to observe trends within and across speakers. On the other hand, when a contrast is weak, due for example to its low functional load, the presence of phonological conditioning (Renwick & Ladd, 2016), or to a near-merger or impending merger (Labov, Karen, & Miller, 1991; Labov, Yaeger, & Steiner, 1972), speakers' intuitions may be unreliable.

Our goal is to test for behavioral trends across a large group of speakers; but, since we hypothesize that speakers will *not* behave uniformly (see 1.2), this work is linked to elicitations of language usage intended to understand variation across geographic space and time. Within sociolinguistics, where the goal is to uncover variation rather than settle on a single norm, data may be collected through written surveys, distributed by mail or

other means (Chambers, 1994, 1998a, 1998b; Nagy, 2001; Schilling, 2014). Many dialectology surveys of linguistic variation have been carried out as interviews in which participants respond to questions about their language use, often by field researchers visiting their home (Kurath & McDavid, 1961; McDavid & O'Cain, 1980; Pederson, McDaniel, & Adams, 1986), with the notable exception of Martinet's (1945) work on French, collected from officers in a World War II prison camp. Specifically for Catalan, Carrera-Sabaté (2006) used single-answer questions to examine the variable realization of pretonic vowels in word-initial position in Western Catalan. She complemented her study with a perception test in which participants "were asked to discriminate possible phonetic differences" (p. 20) and to reflect upon their language use and community norms. The same methods were applied in a study of the alternation between [ɛ] and [a] in compound words (Carrera-Sabaté, 2005).

On the other hand, data for the *Atlas of North American English* (Labov, Ash, & Boberg, 2006), documenting speech variation in the US and Canada, were gathered via recorded telephone calls, while more recent studies, like ours, gather judgments via Internet-based studies (García, 2015; Walker, García, Cortés, & Campbell-Kibler, 2014). The method used here to gather data was in fact more explicitly judgment-based than the aforementioned sociolinguistic studies. We rely on the fact that Catalan speakers are taught the mid vowel distinctions in school, and that they know terms for referring to the high and low mid vowels.

In this way, our study is most similar to questionnaires used to gather pronunciation data for dictionaries, including the *Dictionary of American Regional English* (Cassidy & Hall, 1985). For the first and second editions of his *Longman Pronunciation Dictionary* (1990, 2000), J. C. Wells conducted polls of pronunciation preferences in specific English words, ultimately gathering judgments from nearly 2,000 respondents who provided demographic information on their sex, age, and regional origin along with their "pronunciation preferences" – for example, whether *zebra* is pronounced with a stressed /e/ or /iː/, and whether *scone* rhymes with *cone* or *con* (Wells, 2003, 1999a, 1999b).[2] Countering the possibility that speakers' actual pronunciation may not be reflected in their responses, Wells argues that questions of this type involve phonemic distinctions and thus are well above the threshold of typical linguistic awareness, unlike allophonic aspects of English such as aspiration of stops or the prevalence of syllabic nasals. In a descriptive statistical analysis, Wells shows generational differences that he argues illustrate language change in progress. A similar questionnaire focused on American English was prepared and analyzed by Shitara (1993), highlighting variation within that variety but also with respect to British English.

1.2 Distribution of mid vowels in Catalan

Central Catalan has seven stressed vowels: /i e ɛ a ɔ o u/. In unstressed position, several contrasts are neutralized: /e ɛ a/ surface as [ə], while /o ɔ u/ are realized as [u], and /i/ remains [i] (Mascaró, 2002). The phonological nature of the vowels under study (/e ɛ/ and /o ɔ/) is evidenced by the existence of minimal pairs involving these contrasts: *venen* [ˈbe.nən] 'they come' vs. *venen* [ˈbɛ.nən] 'they sell'; *os* [ˈos] 'bear' vs. *os* [ˈɔs] 'bone'. However, despite the high mid and low mid vowels being contrastive, Mascaró (1984,

2008, 2010) and Wheeler (2005) noticed that their distribution seems to be partially phonologically conditioned.

Based on alternations like *modest* [muˈðest] 'modest' ~ *modèstia* [muˈðɛstjə] 'modesty'; *carbó* [kərˈβo] 'carbon' ~ *carbònic* [kərˈβɔnik] 'carbonic', Mascaró (1984) postulated that prestressing suffixes regularly open the stressed mid vowel in the root, but not other stressed vowels (*sulfur* [sulˈfur] 'sulphur' and *sulfúric* [sulˈfurik] 'sulphuric')[3]. Given that many of the prestressing suffixes in Mascaró (1984) contained a posttonic high front vocoid, Wheeler (2005) examined whether it was the high vowel in the posttonic syllable that conditioned the realization of the stressed vowel. He concluded that there is a "virtually categorical" neutralization of mid vowels (p.44) before posttonic [i, j], such that only low mid vowels can appear before them (*geni* /ˈʒɛ.ni/ 'genius', *gremi* /ˈgrɛ.mi/ 'guild'; *odi* /ˈɔ.ði/ 'hate', *Jordi* /ˈʒɔr.ði/ 'George'). He linked this to a much more subtle tendency for low mid vowels to be favored before vowels /u, o/ in words with penultimate or antepenultimate stress (he presented many counterexamples to that tendency, though). Mascaró (2008) provided more evidence to support both of Wheeler's (2005, p. 44) observations.

Apart from those conditioning factors (i.e., presence of a prestressing suffix or of a posttonic [i] or [j]), immediate context also plays a role in the distribution of the front mid vowels. This is due to a quite complex historical evolution during which the front mid vowels exchanged positions (Rasico, 1993). To simplify, we can say that Latin Ĭ, Ē > Vulgar Latin /e/ > Central Catalan /ɛ/ (but Spanish /e/) and Latin Ĕ > Vulgar Latin /ɛ/ > Central Catalan /e/ (Spanish /i̯e/) (Alkire & Rosen, 2010). The change from Vulgar Latin /e/ to Catalan /ɛ/ (involving an intermediate step /ə/, still preserved in Majorcan Catalan) was quite systematic, whereas the change from /ɛ/ to /e/ was blocked in some contexts (Rasico, 1981). Based on these historical developments, Wheeler (2005) noted that, synchronically, /e/ cannot appear before /l/ (Lat. VĒLA: Cat. *vela* [ˈbɛ.lə], Sp. *vela* [ˈbe.la] 'sail'; but also Lat. MĔL: Cat. *mel* [ˈmɛl], Sp. *miel* [ˈmi̯el] 'honey'), /r/ plus vowel (Lat. CĬRRU: Cat. *cerro* [ˈsɛ.ru], Sp. *cerro* [ˈθe.ro] '(fiber) ball'; but also Lat. TĔRRA: Cat. *Terra* [ˈtɛ.rə], Sp. *Tierra* [ˈti̯e.ra] 'Earth'), or /r/ followed by a coronal consonant (Lat. VĬRĬDE: Cat. *verd* [ˈbɛrt], Sp. *verde* [ˈbeɾ.ðe] 'green, m.'; but also Lat. CĔRTU: Cat. *cert* [ˈsɛrt], Sp. *cierto* [ˈθi̯eɾ.to] 'true, m.'). He added that /e/ is also avoided before [ŋ], /ɲ/, /f/, and /sk/.

Whilst all the contexts mentioned so far condition the presence of a low mid vowel, there is one other context described by Wheeler (2005, p. 41) that raises the front mid vowel. Specifically, a palatal lateral tends to raise the immediately following front mid vowel (*llebre* [ˈʎe.βrə] 'hare', from Lat. LĔPŎRĔ, cf. Sp. *liebre* [ˈli̯e.βɾe]), although Wheeler noted a couple of exceptions. The contexts that favor either high mid or low mid vowels are summarized in Table 1.

Table 1. Conditioning contexts for Catalan mid vowels.

Prestressing suffix	Posttonic [i, j]	Before /r, l/	Before [ŋ], /ɲ, f, sk/	After [ʎ]
/ɛ, ɔ/	/ɛ, ɔ/	/ɛ/	/ɛ/	/e/

Despite the presence of these patterns, other authors (Badia i Margarit, 1968, 1969a, 1969b, 1970a, 1970b; Fabra, 1906; Recasens, 1993) have emphasized the lack of stability of Catalan mid vowels. Badia i Margarit (1969a, 1970b) and Recasens (1993) give lists of words that show alternation between high mid and low mid vowels in the same dialect (Central Catalan) across and/or within speakers. These lists were compiled by the authors through observation of spontaneous oral speech and texts. Regarding the latter case, Badia i Margarit (Badia i Margarit, 1968, 1970a) studied the rhymes of 19th and 20th century poems and concluded that those between non-identical mid vowels (e.g. /e/ with /ɛ/) were common. In fact, imperfect rhymes are also attested in medieval poetry (Coromines, 1974, p. 219) and traditional (Badia i Margarit, 1969b) and contemporary songs.

Recasens (1993) also compiled an extensive list of words and suffixes that contain either low or high mid vowels depending on the speaker, and claimed that two normative pronunciations have to be accepted for some words, rather than just one.

While the sources of this inter- and intraspeaker variation have not been systematically studied, Recasens (1993) ventured that they might include analogy, Spanish influence, or the existence of stress-induced vowel neutralizations. Extensive and continued contact with Spanish, which has only one mid vowel height, might be eroding the mid vowel contrasts in Catalan. In addition, he posited that the fact that some infinitives (e.g. *arrencar* [ərəŋˈka] 'to pluck', *galopar* [gəluˈpa] 'to gallop') do not reveal the identity of the vowel in root-stressed forms may result in unclear intuitions on the part of the speaker (Badia i Margarit, 1970b). According to Badia i Margarit (1970b), another possible source of confusion is the alternation of low and mid high vowels in words that are morphologically related (*pes* [ˈpɛs] 'weight' vs. *peso* [ˈpezu] '(I) weigh'). Finally, the existence of important differences in the distribution of mid vowels across Catalan dialects has been proposed as another potential source of variation (Mora & Nadeu, 2012; Wheeler, 2005).

1.3 Research questions

The main goal of this study is to investigate the phonological intuitions of native Central Catalan speakers regarding the distribution of high mid and low mid vowels. More specifically, we seek to evaluate generalizations about the almost categorical distribution of mid vowels in certain contexts as well as observational data that suggests a high degree of variability for certain words. We do so by testing words with the relevant phonological contexts as well as words extracted from Recasens' (1993) and Badia i Margarit's (1969a, 1970b) lists of variable words. We predict that our participants will exhibit more agreement and more certainty in their choices in words whose mid vowels are phonologically conditioned than in those described as having variable mid vowel height.

Previous claims indicate that younger speakers in urban areas tend to use high mid vowels where older speakers produce low mid vowels (Recasens, 1993); or that, even within the same city, speakers vary in their percentages of use of the mid vowels (Lleó, Benet, & Cortés, 2007); and many studies demonstrate that Spanish-dominant speakers have difficulties with the Catalan mid vowel contrasts (e.g. Bosch et al., 2000; Pallier et

al., 1997). This study thus also sets out to test how much interspeaker variation exists (i.e., to what extent speakers disagree on the mid vowel height of a given word or set of words), and to test the strength of speaker intuitions about mid vowel height for most words. We predict that interspeaker variation will occur, and that most intuitions will be strong. In addition, we explore the role of dominance in Catalan, as self-assessed through a bilingual language history questionnaire, on vowel height choice and certainty of judgment. In particular, we test whether age and dominance in Catalan correlate with rate of low mid vowel choice. We also test whether age and dominance in Catalan correlate with vowel height certainty. Our results will extend and expand upon previous work examining speaker-specific correspondences between intuitions and acoustics (Nadeu & Renwick, 2016), both through comparison to the intuitions gathered in that work, and by testing whether words that exhibit greater intraspeaker variation are also more acoustically variable.

2. Methods

2.1 Materials and procedure

Our stimuli were grouped into three categories of words. Two categories were based on the previous literature (see section 1.2): "Nonvariable" and "Variable" words. In Nonvariable words the phonological context is said to condition mid vowel height (in all cases but one, the context, immediate or not, is expected to trigger a low mid vowel). The Variable words we tested were extracted from published lists of words having variable realizations (allowing both low and high mid vowels) in the same dialectal area or even as produced by the same speaker (Badia i Margarit, 1969a, 1970b; Recasens, 1993). The third category was comprised of words tested in a previous experiment (Nadeu & Renwick, 2016) that paired speakers' intuitions with their own productions (we refer to these words as "Tested" words) in order to start exploring inter- and intraspeaker variability in mid vowel realization.

We selected a total of 220 words: 40 words × 2 conditions (Variable, Nonvariable) × 2 mid vowel sets (front, back), plus 30 Tested words × 2 sets. All words, including their condition within the study, are listed in Appendix A. Table 2 shows the distribution of Nonvariable words by subcondition. One word with a stressed back vowel (*coreògraf* 'coreographer') from the Nonvariable condition was excluded from all analyses due to experimenter error in the way it was presented in the survey. In addition, *llengua* 'language' was accidentally included twice, as Nonvariable (due to the immediate context, see section 1.2) and as Variable (as mentioned in several sources). It is ultimately analyzed as Nonvariable, based on its treatment by participants. The total number of words analyzed is thus 218. Also note that a Nonvariable word with a front mid vowel (*nervi* 'nerve' with a posttonic /i/) was mentioned as a Variable word in one of the sources discussed in section 1.2 (Badia i Margarit, 1970b).

Table 2. Distribution of Nonvariable words by conditioning factor, separately by front and back vowels.

Conditioning Factor	Front	Back
Immediate context	20	NA[4]
Posttonic [i, j]	15	20
Prestressing suffix	0[5]	13
Posttonic [i, j] + Prestressing suffix	5	6
Total	40	39

The words belonging to the category Tested constituted a subset of the words included in Nadeu & Renwick (2016). In that study, stimuli were chosen so that they met certain requirements in terms of word length, syllable structure, immediate phonological context (for ease of segmentation), and following vowel, but no attention was paid to whether words had been identified as variable previously or whether they exhibited one of the vowel lowering contexts. Among the subset included in the current study, 13 Tested words (out of 60) also fell in other categories (words with a front vowel: 5 Nonvariable, 5 Variable words; words with a back vowel: 3 Variable words). In the analyses, we indicate when these 13 words are included in the Variable or Nonvariable conditions and when they are considered as part of the Tested condition. The remaining 47 Tested words (see Appendix A) have no known phonological trigger of vowel quality, nor are they reportedly variable in the literature, so we have no prediction regarding the height of their stressed mid vowels.

For each word, a unique, meaningful sentence was created so that the meaning of the target word was clear from the context. Note that Catalan orthography does not always provide information on mid vowel height: /e/ and /ɛ/ are both represented by <e>; /o/ and /ɔ/, by <o>. When a word requires a stress mark following the Catalan spelling conventions, a different diacritic is used for the open mid and the close mid vowels (/e/ = <é> and /o/ = <ó>; /ɛ/ = <è> and /ɔ/ = <ò>). However, most words do not carry stress marks. In order to reduce the influence of orthography on participants' judgments, in words with orthographical stress marks (which indicate prescriptive mid vowel height), stress marks were removed.[6] The stressed vowel in all target words was bolded and underlined.

Two survey lists were created, each containing 120 words (40 per condition). The two lists only overlapped minimally, with ten words belonging to the Tested category repeated. This allowed us to survey a wide number of the 210 target items studied previously (Nadeu & Renwick, 2016), while maintaining a smaller number of 20 total lexical items judged across the widest possible number of speakers, including those in the previous study. All the other words were only presented in one list or the other. The surveys were administered online using Qualtrics, which presented the sentences to each participant in a randomized order. Beneath each sentence, a scale with the following six responses was presented to collect speakers' judgments: "Clearly Open", "Maybe Open", "Open or Close", "Maybe Close", "Clearly Close", "I don't know this word". The possible responses were listed from left to right. In total, four surveys were created, using each of the two word lists in the two possible orders of responses (from left to right: from "Clearly Open" to "Clearly Close", and from "Clearly Close" to "Clearly Open"), to control for a possible response bias derived from the disposition of the different categories on the screen.

Following this section, participants answered a series of questions about the mid vowel contrasts (whether they think they produce a contrast between high mid and low mid vowels, whether they find it easy to determine what mid vowel height a word has, and whether they are aware that variation in mid vowel use across speakers exists; see Appendix B). Given that no adult monolinguals of Catalan exist, participants also completed the Bilingual Language Profile questionnaire (BLP) (Birdsong, Gertken, & Amengual, 2012), which includes questions about participants' (self-assessed) language history, use, proficiency, and attitudes. The output score can range from −218 (monolingualism in Spanish) to +218 (monolingualism in Catalan), where 0 would represent perfectly balanced bilingualism. We report participants' BLP scores in section 3.1.

2.2 Participants

Participants were recruited through ads in social media, emails sent to several Catalan universities, and word of mouth. A link to the web-based survey was included in the emails and ads that were distributed. Any person 18 years of age or older was allowed to complete the survey.

We obtained 242 complete responses, collected over approximately one month. Of these, we selected for analysis those from speakers of Central Catalan who were born and had lived most of their lives in Central Catalan-speaking areas (n = 146). We decided to exclude speakers of other dialects of Catalan because there were not enough responses per dialect to draw any meaningful comparisons and it is known that the distribution of mid vowels varies importantly across dialects. It is worth noting, however, that, even within the Central Catalan-speaking areas, variation concerning the mid vowels is found, such that one or both contrasts are merged in some places (Recasens, 2014; Recasens & Espinosa, 2009). Some of our participants were from Girona, where the contrast between /o/ and /ɔ/ is neutralized.

The mean age of the 146 speakers finally included in the analyses is 30.31 (SD = 10.5; range = 18–72). The survey took approximately 30 minutes to complete (M = 45.8; SD = 153.7; range = 9.8–1749.7).[7]

2.3 Data processing and analysis

We first ensured the validity of our data by analyzing the BLP score data in relation to participants' responses to metalinguistic questions about the mid vowel contrasts, using Pearson's product moment correlations.

The effect of condition on vowel height judgments was explored by focusing on the distribution of responses by condition. We conducted tests to determine whether it is the case that Nonvariable words are more consistently judged to have a certain vowel across speakers. Fleiss' Kappa tests (Fleiss, 1971) were used to evaluate the level of agreement among survey respondents for each word. Labels for each judgment were created by collapsing the judgment categories of "Clearly Open" and "Maybe Open" into a single "Open" category; likewise for "Close" judgments. A category "Other" combined the "Close or Open" and "I don't know the word" options. Labels were collapsed in this way

primarily for two reasons. First, our possible responses combine vowel height and certainty. By collapsing across "Clearly" and "Maybe" responses we are limiting the analyses to vowel height independently of degree of certainty. Second, "Maybe" responses were very scarce. A by-word summary of participants' responses, which uses the categories "Close", "Open", and "Other", appears in Appendix C.

Responses were coded for each word's condition (Nonvariable, Variable, No Prediction) and vowel set (Back vs. Front). The No Prediction condition includes the 47 Tested words that were not also coded as Variable or Nonvariable. Data from all three conditions were compiled into a single data set, and the test was run across each list (List 1 vs. List 2) separately, using the flag "detail = TRUE" in R's *kappam.fleiss()* command to gain separate scores for each condition and response type across both front and back vowels.[8] In addition, the factors that affected probability of close vs. open responses were explored through a logistic mixed-effects regression, which was run, like all the other tests, in R (R Core Team, 2000), using the *glmer()* function of the lme4 package (Bates & Maechler, 2009, p. 4).

Degree of judgment certainty (whether speakers selected "Clearly" vs. "Maybe" responses) was also modeled using logistic mixed-effects regression. In addition, the relationship between speakers' judgments (height distinctions and degree of certainty) and several factors was explored through correlations.

Finally, the data obtained in the present study were examined in light of the results from Nadeu & Renwick (2016). As noted in the introduction, a subset of the stimuli (Tested words) was taken from that earlier study. We compared the judgments elicited by those stimuli in the present study, against mismatches between participants' intuitions of mid vowel height and their actual phonetic realizations observed in the previous study. The mismatches were obtained through k-means clustering, which assigns individual data points to one of a specified number of clusters (here, two) based on their characteristics (in this case, normalized F1 and F2 values). We applied k-means clustering separately to front and back mid vowels. Using the classification yielded by the k-means clustering, we calculated the number of times that the acoustics of a token did not match the participant's intuition regarding phonological vowel height (see Nadeu & Renwick [2016] for more information).

3. Results

3.1 Participant characteristics

As noted in section 2.2, only responses from participants who were born and raised in Central Catalan-speaking areas were considered (n = 146). The selected speakers still represented a variety of geographical locations, ages, and degree of Catalan dominance. Language dominance was assessed by means of the BLP questionnaire. Our participants obtained a mean BLP score of 69.06 (SD = 47.74; range: −72.01–148.80). These values represent an extensive range of Catalan-Spanish bilingualism, encompassing both Spanish- and Catalan-dominant individuals as well as balanced bilinguals. Figure 1

shows the distribution of BLP scores, which reveals that Catalan-dominant speakers (BLP > 0) outnumber Spanish-dominant ones among the participants.

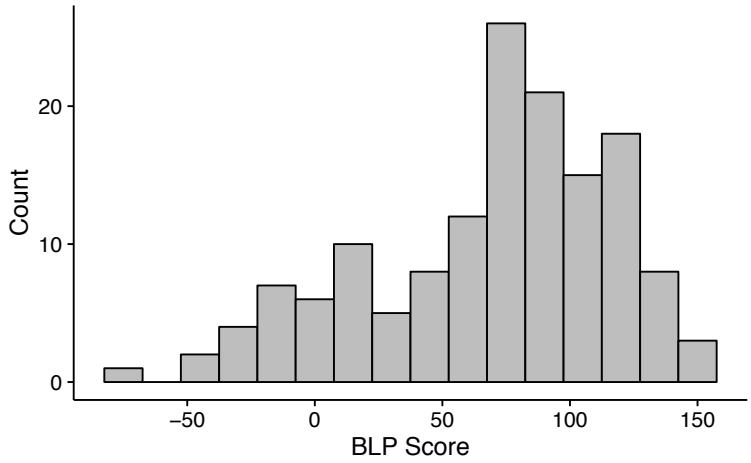


Figure 1. Histogram of the BLP scores of the 146 participants (positive values indicate dominance in Catalan).

Between the vowel judgment survey and the BLP questionnaire, participants answered a few questions about the mid vowel contrasts. Participants indicated how confident they were that they produced the /e – ɛ/ and /o – ɔ/ distinctions in two separate questions (1 = "I believe I do not distinguish them", 10 = "I am sure I distinguish them"), how easy it was for them to determine whether a word had a low mid vs. high mid vowel (1 = "Very difficult", 10 = "Very easy"), and whether they had noticed that speakers vary in their mid vowel use (1 = "I have never noticed that", 10 = "I have noticed that many times"). The scatterplots in Fig. 2 show the distribution of the answers to these four questions.

Fig. 2 displays participants' self-reported degree of confidence that they produced the mid vowels distinctly ("Back Mid Vowel Distinction": mean = 8.75, SD = 1.87; "Front Mid Vowel Distinction": mean = 8.89, SD = 1.71). The scatterplots show a concentration of answers toward the upper end of the scale, indicating that, in general, participants are confident that they produce the mid vowel contrasts. Values below 5, indicating that speakers were not so sure or less confident that they produced the distinction, are scarce, and the lowest value in the scale (i.e., 1 = "I believe I do not distinguish them") was not selected by any participant. In addition, the upper right scatterplot reveals that most speakers can easily determine which mid vowel a word has, again with most values concentrating in the higher end of the scale (mean = 8.47, SD = 2.08). These responses are an assurance that participants are capable of providing judgments, which yield the bulk of the data analyzed in the following subsections. Finally, the lower right scatterplot shows that most participants are aware of variation in the distribution of mid vowels, such that two speakers may disagree in which mid vowel they assign to the very same word (mean = 9.02, SD = 1.82). Very few speakers reported not being aware of such variation. This shows that the phenomenon under study is salient enough that linguistically naïve speakers are aware of it.

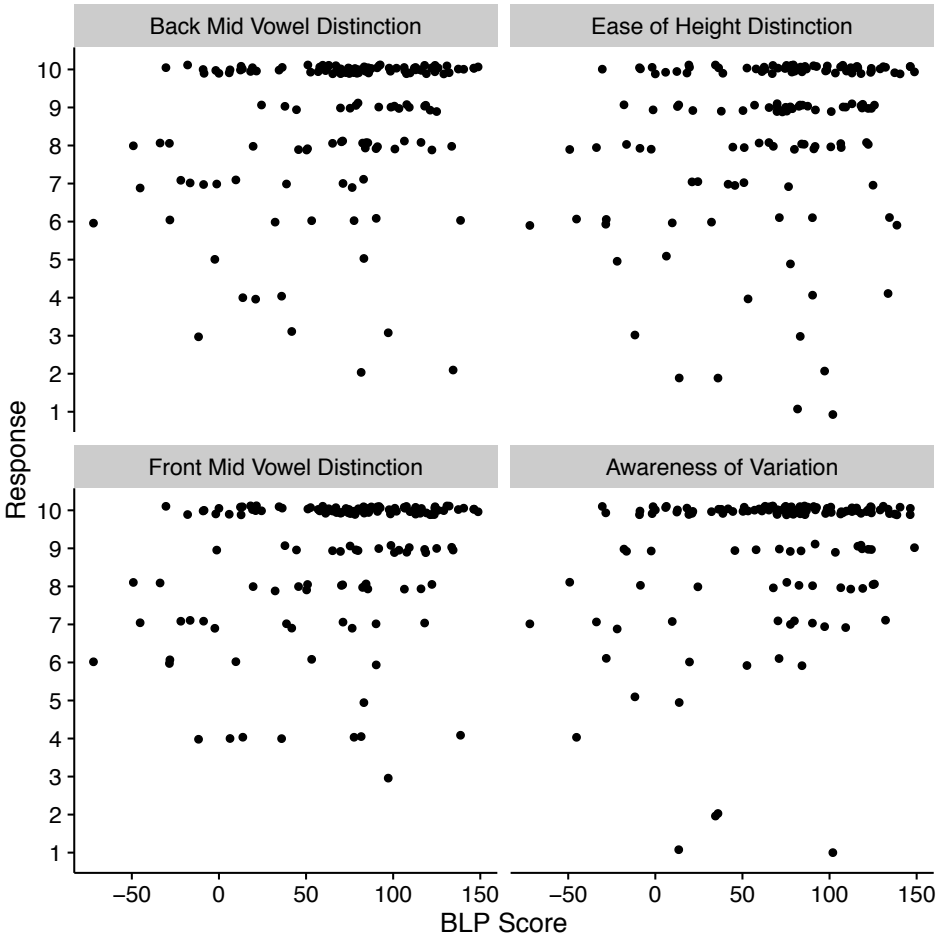


Figure 2. Scatterplots of responses (ranging from 1 to 10) to each of the four questions gauging participants' awareness of mid vowel contrasts. Responses are jittered vertically to avoid overplotting.

The responses to the four questions concerning mid vowel contrasts are positively, although weakly, correlated with BLP scores, as shown in Table 3. As BLP scores increase (indicating higher dominance in Catalan), participants feel more confident that they can distinguish between low mid and high mid vowels. Participants with higher BLP scores also report finding it easier to decide which mid vowel a word has and having noticed variation more frequently.

Table 3. Results of the Pearson's product moment correlation between BLP scores and responses to the four questions aimed at gauging participants' awareness of mid vowel contrasts.

Comparison	*r*	*p*-value
BLP score – Back Mid Vowel Distinction	0.24	< 0.01
BLP score – Front Mid Vowel Distinction	0.28	< 0.001

BLP score – Ease of Height Distinction	0.20	< 0.05
BLP score – Awareness of Variation	0.25	< 0.01

Taken together, the participant characteristics reported here indicate that the judgment data were obtained from speakers with a reasonable level of Catalan dominance and a reasonable level of confidence that they were able to complete the main task of the experiment. Nonetheless, we restrict the remainder of our analysis to judgments provided by the 128 participants who are Catalan-dominant (BLP > 0) speakers of Central Catalan varieties. We take this step because it is known that dominance in Catalan vs. Spanish affects performance on mid-vowel distinctions, and because our research focus is the stability of the mid-vowel contrast within the Catalan-speaking community, not across Catalan- vs. Spanish-dominant populations. In the following subsections, we analyze mid vowel judgments, relating those responses to the participants' characteristics reported here.

3.2 Open vs. close distinction

The total counts of responses from Catalan-dominant participants (n = 15,300) are plotted in Fig. 3, separated by condition and front-back distinction. A cursory observation reveals a clear distinction between words in the Nonvariable category and words in the other two conditions. The No Prediction category includes most of the words that were previously tested in an acoustic study of the mid vowels (see Section 2.1). As noted there, some of those words had either the phonetic context to be included in the Nonvariable group or had been listed as Variable. The No Prediction group excludes those that were recoded as belonging to one of the other two categories.

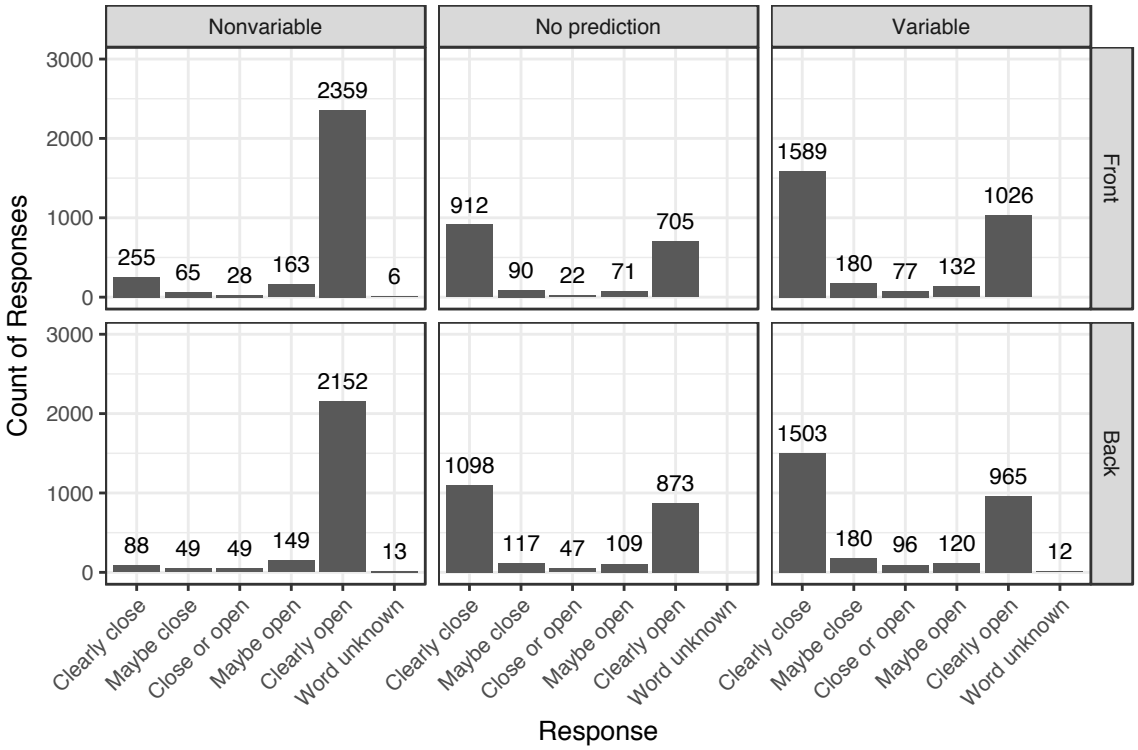


Figure 3. Counts of Catalan-dominant participants' responses by condition, separated by front and back mid vowel sets.

In the Nonvariable category, we note a preponderance of "Clearly Open" answers, as predicted. Recall that the Nonvariable condition consists of words where the phonetic context predicts mid vowel height, such that the low mid vowels are expected. For the back vowels, the second most frequent answer is "Maybe Open". However, for the front vowel set, "Clearly Close" judgments outnumber "Maybe Open" answers. This pattern is due to the fact that that one of the contexts included for the front vowels triggers raising, rather than lowering (cf. Table 1). More specifically, the Nonvariable condition includes five words that begin with the palatal lateral /ʎ/, which, according to Wheeler (2005), has a raising effect on the front mid vowels. Out of the five words included, two (*lleixa* 'shelf'*, lleig* 'ugly') have that predictor only, and "Clearly Close" or "Maybe Close" responses make up over 85% of the judgments that they received.

For the remaining three /ʎ/-initial words, the situation is a bit more complex. In *llengua* 'word, tongue' and *llenya* 'firewood', the consonants following the stressed vowel ([ŋ] and /ɲ/, respectively) trigger vowel lowering, which conflicts with the effect of the preceding consonant. The results show that these two words behave differently. *Llenya* is clearly nonvariable, and the effect of /ʎ/ is obvious (95% of "Clearly Close" and "Maybe Close" answers). *Llengua*, however, behaves like a Variable word (Badia i Margarit, 1969a, 1970b; Recasens, 1993), with slightly more close (66%) than open answers. The fifth word, *llei* 'law', is included in Recasens's (1993) list of variable words. Its behavior, nevertheless, does not show much variability: over 82% "Clearly Close" and "Maybe

Close" judgments. To sum up, the tendency to assign a close /e/ following /ʎ/ is quite strong, as shown in four of the five words that have this context.

Turning to the No Prediction and Variable conditions, the patterns are comparable across the front vs. back dimension. The most frequent answer is "Clearly Close"; and the second most frequent, "Clearly Open" (see Fig. 3), contrasting with the behavior exhibited by the Nonvariable words. In addition, there is not such a clear bias toward one of these responses as in the Nonvariable condition. The similar surface patterns between the No Prediction and Variable conditions mask important differences in the degree of agreement across speakers between these two conditions, which are explored in Section 3.5.

Figure 3 shows that in the vast majority of cases participants selected one or the other height. "Close or Open" responses, indicating that they did not know which mid vowel height to assign or thought the word could be pronounced either way (despite being familiar with the word), are scarce (n=319/15,300; 2.1% of the data). Table 4 shows the counts of words grouped by the percentage of ambiguous ("Close or Open") responses. In fact, only one word (*consell*) received more than 10% ambiguous responses, indicating that participants strongly prefer to select a single, fixed vowel height rather than a variable height.[9]

Table 4. Number of words with ambiguous responses by condition.

Condition	Words with no ambiguous responses	Words with 0-10% ambiguous responses	Words with over 10% ambiguous responses	Total words
Nonvariable	36	48	0	84
No Prediction	23	24	0	47
Variable	19	68	1	88

In sum, participants are generally able to make a judgment one way or another, and tend to refrain from using the "Close or Open" response. Out of 219 words, only 25 (11.4%) elicited more than 5% of ambiguous responses. Ambiguous responses are included in the Fleiss' Kappa tests presented below, but are excluded in the remaining analyses.

To determine the amount of agreement raters obtained in each of the three conditions, two Fleiss' Kappa tests were run: one on each of the two survey lists. Remember that we collapsed "Clearly Open" and "Maybe Open" responses into a single "Open" category. The same was done for the "Close" answers.

The Fleiss-Kappa results (Table 5) show that for Nonvariable words, the highest, almost perfect kappa scores obtain with the label "Open", both for List 1 (kappa = 0.895) and List 2 (kappa = 0.917). For Nonvariable words, the "Unknown" and "Close" labels have very low (though still significant) kappa scores, reflecting the very sporadic distribution of those responses in that classification. One exception is for Nonvariable front "Close", which includes the handful of words conditioned to have a close vowel (*llenya, llei,* etc.). Among the No Prediction words, kappa scores are lower, but still represent substantial (or almost perfect, in one case) agreement. Finally, the Variable words show the lowest kappa scores, with moderate (or substantial) agreement. Responses of "Unknown",

collapsing "Close or Open" and "I don't know the word" answers, are rare and thus have low kappa scores throughout.

Table 5. Results of the two Fleiss' Kappa tests showing agreement across Catalan-dominant participants.

			List 1			List 2		
			Kappa	***z* value**	***p*-value**	**Kappa**	***z* value**	***p*-value**
Nonvariable	Back	Open	0.90	412.53	< 0.001	0.92	479.44	< 0.001
		Close	0.06	26.42	< 0.001	0.06	29.41	< 0.001
		Unknown	0.02	8.72	< 0.001	0.04	20.08	< 0.001
	Front	Open	0.92	423.11	< 0.001	0.91	477.65	< 0.001
		Close	0.61	278.60	< 0.001	0.36	188.91	< 0.001
		Unknown	0.01	2.14	< 0.001	0.02	9.41	< 0.001
No Prediction	Back	Open	0.75	347.66	< 0.001	0.76	396.72	< 0.001
		Close	0.78	360.09	< 0.001	0.82	430.06	< 0.001
		Unknown	0.03	11.97	< 0.001	0.02	11.06	< 0.001
	Front	Open	0.76	348.59	< 0.001	0.81	423.72	< 0.001
		Close	0.83	384.46	< 0.001	0.82	430.48	< 0.001
		Unknown	0.01	5.62	< 0.05	0.01	4.44	< 0.001
Variable	Back	Open	0.58	269.09	< 0.001	0.66	346.42	< 0.001
		Close	0.73	336.34	< 0.001	0.76	397.97	< 0.001
		Unknown	0.06	27.61	< 0.001	0.03	17.13	< 0.001
	Front	Open	0.58	269.22	< 0.001	0.54	283.26	< 0.001
		Close	0.68	313.38	< 0.001	0.71	371.11	< 0.001
		Unknown	0.04	19.93	< 0.001	0.03	15.34	< 0.001

The factors affecting individual judgments of mid vowel height by Catalan-dominant participants were evaluated in a logistic mixed-effects regression, whose output is shown in Table 6. The binomial dependent variable in the model, fit by maximum likelihood, was Response, coded broadly as either "Open" or "Close". The total number of responses considered was 14,950, while the smaller number of responses (350) given as "I don't know the word" or "Close or Open" were not considered due to their sparseness, representing 2.3% of all Catalan-dominant responses. Fixed factors were Vowel Frontness (Front vs. Back); Condition (Nonvariable, Variable, No Prediction); BLP score; Age (years); Sex (Male or Female); List (1 or 2, treated as a factor); and Response Order (left to right: "Close" to "Open", or "Open" to "Close"). Additionally, an interaction between Vowel Frontness and Condition was included. Random intercepts were fit for each participant, and the continuous variables BLP score and Age were rescaled to z-scores.[10] An early model additionally included an interaction of BLP Score with Age, but this was removed due to its lack of significance.

Based on descriptive observation of the data and reports in the phonological literature, it was predicted that an increase in "Open" responses would occur for words in the Nonvariable category, particularly among speakers who had higher ages, and potentially also among speakers with higher BLP scores even within this Catalan-dominant group. This would result in an increase in log odds for "Open" responses under those circumstances. In the model presented here, a positive coefficient indicates an increased

likelihood of a "Close" response, while a negative coefficient indicates an increased likelihood of an "Open" response.

Table 6. Fixed effects from logistic mixed-effects regression evaluating likelihood of "Open" vs. "Close" response among Catalan-dominant speakers. The default levels of the categorical predictors are: Frontness = Front; Condition = Nonvariable; Sex = Female; List = 1; Order = ["Open" to "Close"].

Variable	Coefficient	Std. Error	*z* value	*p* (>|z|)
(Intercept)	−2.06	0.07	−29.11	<0.001
Frontness = Back	−0.76	0.11	−7.11	<0.001
Condition = No Prediction	2.34	0.08	30.52	<0.001
Condition = Variable	2.51	0.07	35.44	<0.001
BLP Score	−0.01	0.03	−0.48	0.632
Age	−0.06	0.03	−2.56	0.011
Sex = Male	0.00	0.06	0.05	0.957
List = 2	−0.04	0.05	−0.71	0.476
Order = ["Close" to "Open"]	−0.01	0.05	−0.20	0.845
Frontness=Back*Condition=No Prediction	0.71	0.12	5.73	<0.001
Frontness=Back*Condition=Variable	0.77	0.12	6.45	<0.001

The model demonstrates that several phonological factors do significantly affect participants' choice of mid vowel height, as do some demographic factors. The main effect of Frontness indicates that when participants are judging a back vowel, they are 0.47 times more likely to give an "Open" response ($e^{-0.755} = 0.47$) in the Nonvariable condition. However, when the vowel was Back, but the Condition was No Prediction or Variable, shown in the interaction terms, participants were more likely to give a "Close" response (log odds increased by 2.04 and 2.15 times, respectively). This is an initial statistical confirmation that participants responded differently to Nonvariable words. The main effects for Condition provide additional support for this prediction: a "Close" response was 10.35 times more likely in the No Prediction condition and 12.26 times more likely in the Variable condition.

Turning to demographic factors, only Age had a significant effect on Response; in this case, an increase in the independent variable resulted in an increase in "Open" responses. For Age ($p < 0.05$), a one-unit increase (one standard deviation, approximately 10.5 years) gave a likelihood increase of 0.94 times for "Open" responses. However, BLP Score was not significant ($p = 0.632$), indicating that among Catalan-dominant speakers, the *degree* of dominance does not affect the likelihood of an "Open" response. Sex was not significant ($p = 0.957$), indicating that men and women responded equivalently. The two survey-specific factors List and Order were also not significant ($p = 0.476$ and $p = 0.845$, respectively); this confirms that participants responded equivalently regardless of list and order.

The results of this logistic regression model confirm the findings of our descriptive analysis and Fleiss-Kappa tests; i.e. that participants responded differently to words in the Nonvariable class than to words in the No Prediction or Variable conditions. The results also indicate an additional source of predictable variation, Age, which suggests that older

speakers are more likely to give "Open" responses than younger speakers. While the effect of BLP Score trends in the same way (its coefficient is negative, indicating an increased likelihood of Open response with an increase in BLP Score), its effect is not statistically significant. Additionally, the interaction of Age with BLP Score, among Catalan-dominant speakers, is not significant. The specific correlations of BLP score and Age to the results were tested for front and back vowels separately (with separate correlations for List 1 and List 2, as they contain different words), but all correlations were found to be insignificant ($p > 0.05$).

3.3 Certain vs. not-certain distinction

While the previous section evaluates responses according to "Open" vs. "Close" decisions, this section considers whether phonological and demographic factors affect participants' certainty of response. It is predicted that Nonvariable words, seen above to elicit more "Open" responses, should elicit greater certainty from participants than No Prediction or Variable words. The effects of BLP Score and Age are also tested.

A logistic mixed-effects regression, again a generalized linear mixed model fit by maximum likelihood, was run using responses from Catalan-dominant speakers, similarly to the model in Table 6. The binomial dependent variable was Certainty, collapsed into "Clearly" and "Maybe" responses; these were created by combining "Clearly Open" and "Clearly Close" into the first category, and "Maybe Open" and "Maybe Close" into the second. Again, responses of "I don't know the word" or "Close or Open" were excluded. Fixed factors were Condition (Nonvariable, Variable, No Prediction); BLP Score; Age; Response (broadly, as above, either "Close" or "Open"); and order of presentation. Random intercepts were fit for each participant, and scaled values of BLP Score and Age were used, as above. This was the most complex model that converged.

In this model, shown in Table 7, an increase in likelihood of "Clearly" responses is associated with a negative coefficient, while an increase in the likelihood of a "Maybe" response is associated with a positive coefficient.

Table 7. Fixed effects from logistic mixed-effects regression evaluating likelihood of "Clearly" vs. "Maybe" response. The default level of Condition is Nonvariable, while the default level of Response is "Open", and the default level of Order is ["Open" to "Close"].

Variable	Coefficient	Std. Error	*z* value	*p*(>|z|)
(Intercept)	−3.95	0.30	−13.31	<0.001
Condition = No Prediction	0.27	0.10	2.81	<0.01
Condition = Variable	0.45	0.09	5.01	<0.001
BLP Score	−0.64	0.22	−2.92	<0.01
Age	−0.32	0.22	−1.40	0.161
Response = "Close"	0.09	0.08	1.21	0.225
Order = ["Close" to "Open"]	−0.55	0.44	−1.25	0.211

As shown in Table 7, Condition does have a significant effect on participants' certainty, as do a range of other factors. The effects for Condition indicate that "Maybe" responses

do increase significantly for the No Prediction and Variable words, with respect to default Nonvariable words: the odds of a "Maybe" response increase 1.31 times for No Prediction words, and 1.56 times for Variable words. BLP Score has a significant effect ($p < 0.01$) in the opposite direction: as Catalan dominance increases by one standard deviation, the likelihood of a "Clearly" response increases 0.53 times. The other demographic factor, Age, is associated with a small increase in "Clearly" responses, but it is not significant ($p = 0.161$). Two other factors were tested, but do not reach significance: the category that participants chose ("Close" vs. "Open") has no significant effect ($p = 0.225$), nor does the order of stimulus presentation from left to right ($p = 0.211$).

The results of this logistic regression model indicate, as predicted, that Nonvariable words are more likely to elicit high judgment certainty from participants, an effect that is increased for individuals with a high BLP score. The significant effect of BLP score on Certainty was tested in further detail. Because the two different word lists, as well as front vs. back vowels, may have elicited different response patterns, four correlations were tested, between BLP and Certainty for front and back vowels and for each list. A positive correlation was found between Certainty and BLP score for front vowels in List 1 ($r = 0.24$; $t(58) = 1.9269$, $p < 0.05$), but all other correlations were insignificant ($p > 0.05$). Thus there is only meager evidence that higher Catalan dominance results in greater certainty of vowel height judgments, within a pre-selected group of Catalan-dominant individuals.

3.4 Certainty vs. intuitions of close-open contrasts

It is hypothesized that speakers who have greater certainty about their vowel judgments in specific words will also be more likely to give high scores to the post-survey questions regarding the mid vowel contrasts. This was evaluated by calculating a speaker-specific certainty score, which was matched to each participant's response to post-survey questions. The certainty score is the number of "Clearly" responses given by each participant, as a percentage of total responses. Scores were calculated separately for front and back vowels for two reasons: first, because some participants are from Girona, where the back height contrast is neutralized (Recasens & Espinosa, 2009); second, because speakers responded separately regarding their intuitions of the front vs. back vowel distinctions. Certainty scores were then correlated with responses to three post-survey questions: participants' confidence of producing an /e – ɛ/ distinction; their confidence of producing an /o – ɔ/ distinction; and their ease of discriminating the open and close vowels.

The results of these are plotted in Fig. 4 and Fig. 5. In both figures, it is clear that high-confidence responses of 10 are very common; additionally, however, participants had high certainty overall, as also shown in 3.1. The mean responses regarding a distinction in pronunciation (Fig. 4) were 9.03 for front vowels (SD 1.65) and 8.90 for back vowels (SD 1.81). The mean response regarding participants' ease of making the distinction (Fig. 5) was 8.56 (SD 2.08). Participants' mean certainty scores were 89.14% for front vowels (SD 20.39%) and 87.65% for back vowels (SD 19.98%). In each case the correlation is strong and positive, as summarized in Table 8. We must note, however, that the

correlation may be partially due to our methods: since questions about the mid vowel contrasts were presented after participants had completed the survey, it is possible that the responses to the questions were conditioned by participants' performance in the survey itself. That is, participants may not have been aware of the difficulty or ease with which they could distinguish between mid vowels before they completed the survey.

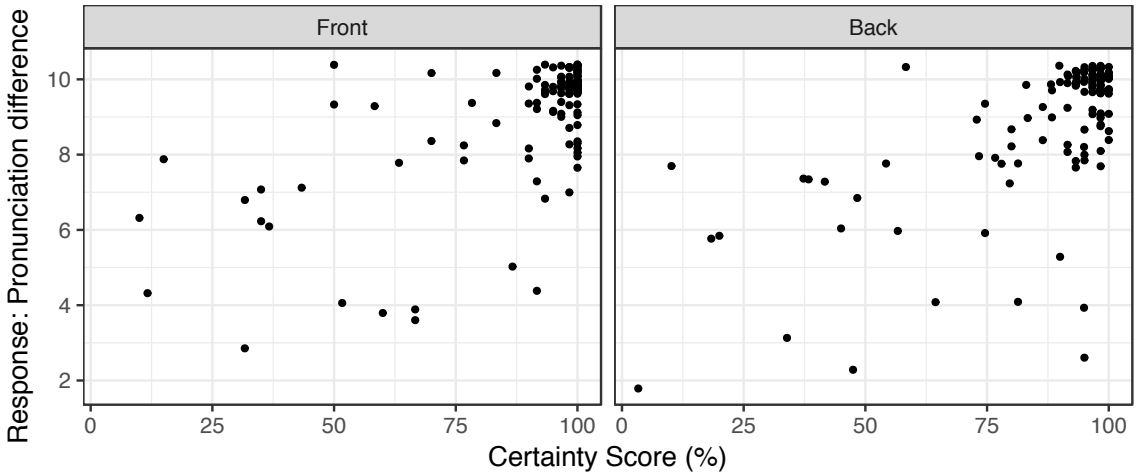


Figure 4. Certainty score based on survey responses vs. responses (ranging from 1 to 10) to questions gauging Catalan-dominant participants' use of mid vowel contrasts. Responses are jittered vertically to avoid overplotting.

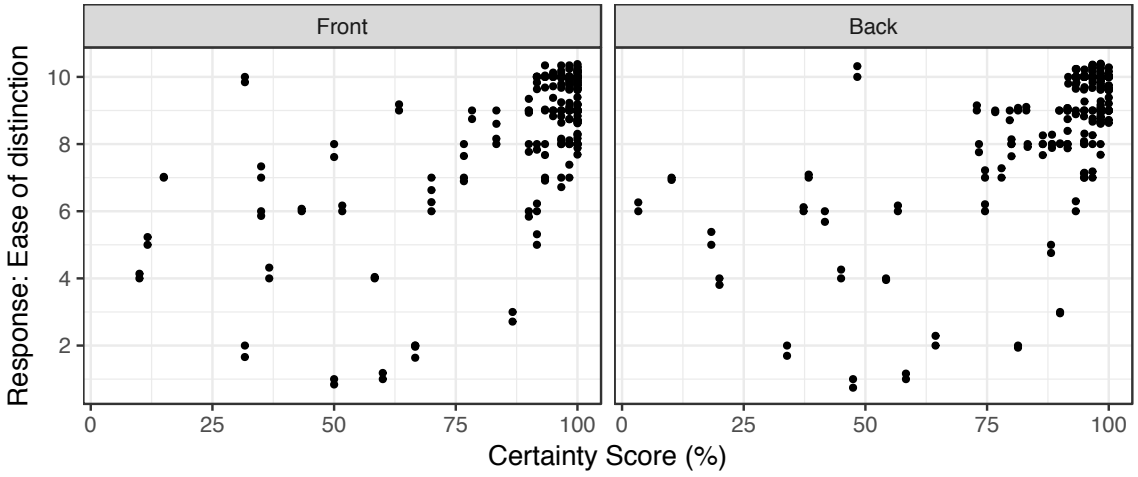


Figure 5. Certainty score based on survey responses vs. responses (ranging from 1 to 10) to questions gauging Catalan-dominant participants' ease of making mid vowel contrasts. Responses are jittered vertically to avoid overplotting.

Table 8. Results of the Pearson's product moment correlation between Certainty score and responses to questions gauging Catalan-dominant participants' use and awareness of mid vowel contrasts.

Comparison	***r***	***p*-value**
Front Certainty score – Front Mid Vowel Distinction	0.66	< 0.0001
Back Certainty score – Back Mid Vowel Distinction	0.66	< 0.0001
Front Certainty score – Ease of Height Distinction	0.66	< 0.0001
Back Certainty score – Ease of Height Distinction	0.67	< 0.0001

3.5 High-agreement and low-agreement words

Figure 3 shows that, particularly among words in the Variable and No Prediction categories, responses of "Open" and "Close" were much more comparable in frequency (though with more "Close" responses) than in the Nonvariable class, where "Open" responses dominate. However, what is not revealed by Fig. 3 is whether participants responded consistently to individual words in the Variable and No Prediction classes, or whether there was a mixture of responses for each word. If words classified as Variable deserve that label, they should show a mixture of "Open" and "Close" responses, for example. To illustrate word-level patterns, the number of responses of each of four types ("Clearly Open", "Maybe Open", "Maybe Close", "Clearly Close") were transformed to percentages of the total number of responses, and plotted in Fig. 6. For compactness and readability, Fig. 6 includes only words with 85% or less agreement by Catalan-dominant participants: thus 38 front-vowel words and 28 back-vowel words are included, while the remaining 152 words are not shown. For each word, four percentages are represented: "Clearly Open", "Maybe Open", "Maybe Close", "Clearly Close". Because the "Maybe" responses were found to be much more rare than "Clearly" responses, they are shown as lines rather than points, for visual clarity near the bottoms of the plots. The x-axis in both plots (Front and Back) is sorted by the "Clearly Close" measure (sorting by "Clearly Open" would produce similar results). Each number along the x-axis represents a stimulus (see Appendix A for mapping), and point shape is determined by phonological condition (Nonvariable, No Prediction, Variable).

The results at the edges of Fig. 6 approach zones of high agreement for both "Open" and "Close" vowel heights, where the majority of tested words fall (>85% agreement), with largely complementary percentages of "Clearly Open" vs. "Clearly Close". However for both Front and Back vowels, there is a core of words that have lower, or mixed, agreement on vowel height. These words, near the transition between high percentage of "Clearly Close" and high percentage of "Clearly Open", range between 40% and 75% agreement across participants. In other words, at least one in four participants would have a different judgment from the others. For some words, agreement is split nearly 50 – 50 among "Clearly Close" and "Clearly Open" responses. These words therefore elicit variable responses, and indeed, the Variable stimuli (square points) are heavily represented in this low-agreement zone, being far more common there than Nonvariable or No Prediction words.

This finding additionally illuminates a difference between the Variable and No Prediction results in Fig. 3: the apparently similar results are actually due to different patterns. In the Variable condition, the same word may have different ratings, whereas in the No Prediction condition, words tend to be labeled with one height category or the other. This is reflected in the presence of many Variable words in the "transition zone" where the percentage of "Clearly Open" and percentage of "Clearly Close" points cross one another, without many words from other conditions.

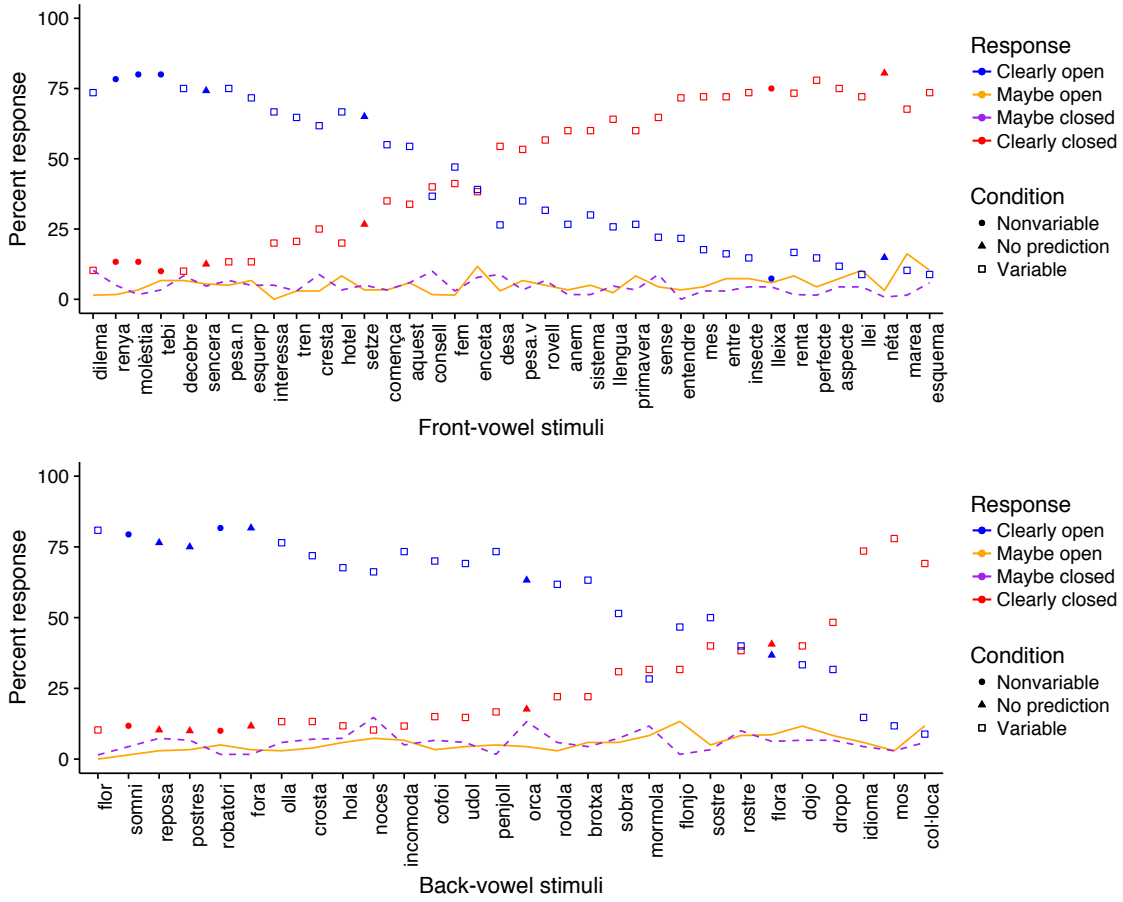


Figure 6: Percent response for four response types, per stimulus. For visualization, words shown include only those with 85% agreement or less by Catalan-dominant speakers. Data are sorted from lowest percentage of "Clearly Closed" answers (left side) to highest. "Maybe Open" and "Maybe Closed" are plotted as lines only.

3.6 Comparison with earlier acoustic study

The set of Tested words represent a subset of the stimuli from a previous study (Nadeu & Renwick, 2016) that compared speakers' pronunciations of mid vowels to their intuitions of height in the same words. We now compare the intuitions and pronunciations acquired

from those 14 speakers to the survey data presented here. First, we examine whether participants in the two studies gave similar judgments. For each word, we calculated the percentage of "Open" survey responses ("Percent Open") in each study. For the survey data, this involved combining "Clearly Open" and "Maybe Open" responses, while the previous study only allowed "Open" and "Close" (or "I don't know") as responses. When percentage of "Open" responses is high, its inverse, percentage of "Close" responses, is low; however, when participants' survey responses had low agreement, percentage of "Open" responses is intermediate (shown above in 3.5). Sixty words were included in the Tested condition, but they could be mapped onto only 58 words in the previous study. Our stimuli included the orthographic words *pesa* 'weight / (s/he) weighs (v)' and *pesca* 'fishing (n) / (s/he) fishes (v)', which can both be either a noun or a verb. In the previous study, we did not specify which part of speech we intended; but in the present survey, we elicited both noun and verb explicitly (four separate lexical items). To compare the results, we mapped the words *pesa* and *pesca* from the acoustic study onto the noun judgments in the survey results, because most words in the previous study were unambiguously nouns, so we assume that speakers might have treated *pesa* and *pesca* similarly.

Figure 7 shows the percentage of "Open" results for each word, across the two studies, plotted against one another. It is clear that the results of the two studies correlate strongly with one another; the correlation was very high for front vowels ($r = 0.927$, $t(26) = 12.63$, $p < 0.001$) and back vowels ($r = 0.928$, $t(28) = 13.18$, $p < 0.001$). Thus we can conclude that participants across the two studies share intuitions for most specific lexical items tested here. There are a few intermediate points: among front vowels, three words fall in between the two clusters; these are *setze* 'sixteen' (No Prediction; Survey: 70.00%; Acoustic: 4.88%), *cresta* 'crest' (Variable; Survey: 70.59%; Acoustic: 26.83%), and *enceta* '(s/he) starts (v)' (Variable; Survey: 46.88%; Acoustic: 63.41%). Among back vowels, two words fall in neither cluster: *sostre* 'ceiling' (Variable; Survey: 53.33%; Acoustic: 7.32%) and *flora* 'flora' (No prediction; Survey: 42.97%; Acoustic: 51.22%). Most of these words are expected to be variable, though in two cases (*setze*, *sostre*) the Acoustic results show a strong preference for "Close" responses. We have no clear explanation for the divergence in judgments on these two words across the two studies, but note that the Acoustic study had a much smaller sample of speakers (n = 14) and that the participants in the Acoustic study were all born and raised in the same city.

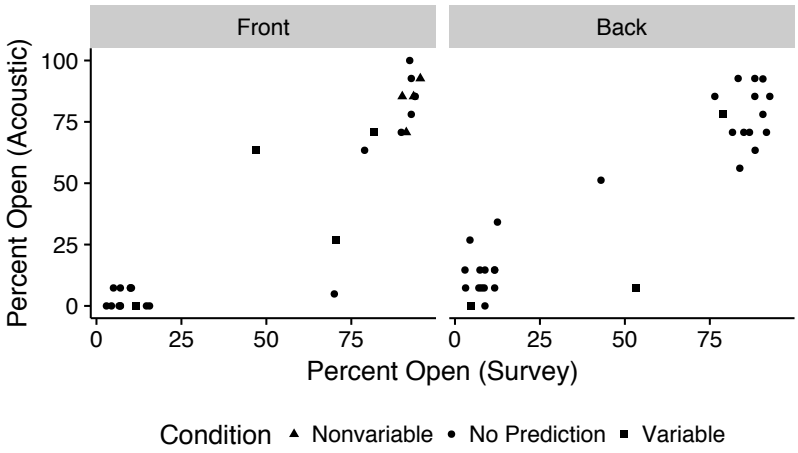


Figure 7. Percentage of "Open" responses across two studies: survey data from Catalan-dominant speakers (Survey) vs. earlier study (Acoustic). Each data point represents one word, its shape coded by Condition.

In the acoustic portion of the earlier study, speakers repeated each stimulus three times, and k-means clustering was used to evaluate whether their productions were consistent and matched their intuitions. When the cluster assigned by k-means did not match the height category assigned by the speaker, a *mismatch* occurred; mismatches were counted across all speakers, for each word. For comparison to the present study, it was hypothesized that words with an intermediate percentage of "Open" responses (a real mix of "Open" and "Close" responses) would, in the acoustic study, be more likely to show a higher number of mismatches across speakers.

Indeed, this comparison found that survey words that have the lowest agreements and are maximally variable correspond to those with highest numbers of mismatches. This is shown in Fig. 8 below. Words that have either low or high percentage of "Open" responses, either below 25% or above 75% on the vertical axis, strongly tend to have a low number of mismatches in the acoustic results – typically fewer than 15 across all speakers (out of 41 data points per word). However, words with low survey agreement, between approximately 25% and 75% on the vertical axis, do tend to have a higher number of mismatches (between 17 and 39 across speakers), affecting both front and back vowels. These low agreement words, echoing the result in Fig. 6, are few in number; but it is clear that the degree of response mixture is highest in words that also have high mismatches. This indicates that variation does not only occur between speakers, but also within speakers.

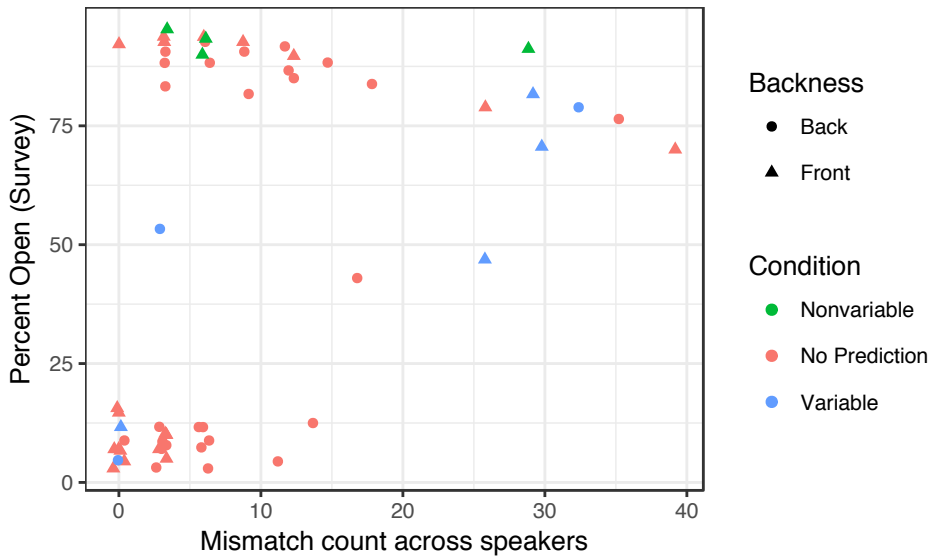


Figure 8. Percentage of "Open" response score based on survey responses, vs. mismatch count from previous acoustic study, per word. Responses are jittered to avoid overplotting.

4. Discussion

4.1 Summary

Participants were asked to identify the height of a stressed mid vowel in a series of words embedded in meaningful sentences. They tended to select "Clearly" responses more often than any other category and only very rarely selected "Either Open or Close" responses, which suggests that native speakers have clear intuitions for most words and that they are able to reflect metalinguistically on them. This is not surprising, given the fact that Catalan speakers learn to do this task in school. Based on the ages of our participants, most of them would have learned to do it. In addition, the fact that they were able to complete the task indicates that they were familiar with the terms used to distinguish the two mid vowel heights. More support for the reliability of speakers' intuitions comes from questionnaire items in which the vast majority of participants reported being very confident that they produced the mid vowel distinctions, as well as ease in distinguishing between close and open mid vowels. In addition, most speakers indicated they were well aware of the existence of variation in mid vowel use across speakers. In other words, the type of interspeaker variation that we analyze is obvious to most native speakers. Our analyses showed that the answers to these survey questions correlate with BLP scores: the more Catalan dominant a speaker is, the higher their degree of confidence that they produce mid vowel distinctions, the easier they find it to distinguish close and open mid vowels, and the more aware they are of the existence of variation.

Turning to the intuitions themselves, our analyses of responses by Catalan-dominant speakers reveal that words in the Nonvariable condition show more agreement across

speakers: the most common answer, by far, is “Clearly Open”. We can thus confirm that the distribution of mid vowels in Catalan is partially conditioned by the phonological context, as proposed by Wheeler (2005) and Mascaró (1984, 2008, 2010). More specifically, certain immediate (e.g. a following [l]) and non-immediate (e.g. the presence of a posttonic [i] or [j]) phonological contexts co-occur frequently with mid vowel lowering. While Mascaró (1984: 124) indicates that the effect of prestressing suffix is active, our data do not allow us to offer any insight into this matter. That is, it is unclear, based on our data, whether the phonological conditioning that we observe is synchronically active or not. We also show that words listed elsewhere as variable and words for which no predictions were made do not favor one height over the other, and they elicit lower degrees of agreement across speakers. Therefore, our results also support impressionistic claims by Badia i Margarit (1969a, 1970b) and Recasens (1993) that some words do not have a fixed vowel height, as evidenced by inter- and intraspeaker variation within the same dialect. However, within our Variable condition, some words are more variable than others, which was not predicted based on this previous literature. Since some of the lists we consulted were based on non-systematic data collection, it is possible that variability was over- or underreported for some words. Furthermore, some of those lists were generated approximately 50 years ago. The variability of some of these words in the Catalan-speaking population as a whole may have increased or decreased since then. Unfortunately, we are not able to test these possible explanations using the present data.

Age, but not language dominance, was found to affect the likelihood of “Open” vs. “Close” answers. Older speakers are more likely to identify vowels as being open.

Participants’ responses to our online survey did not only encode vowel height choice, but also degree of certainty. We thus analyzed whether degree of certainty is mediated by word condition and any demographic factors. The statistical analyses reveal that the No Prediction and Variable conditions are associated with an increase in “Maybe” answers (i.e., less certainty). In addition, language dominance has an effect: the higher the dominance in Catalan, the higher the degree of certainty speakers expressed. As expected, certainty also correlated with participants’ confidence that they produced the mid vowel distinctions and with the ease with which they distinguished open from close mid vowels.

We also plotted the degree of agreement across participants by word. While most words exhibit high agreement (over 75% of speakers chose the same vowel height), some do not. Among those that do not, words in the Variable condition are overrepresented. This indicates that words in the Variable condition tend to be the most variable words in the survey.

Finally, we compared the degree of agreement of words in the Tested category to mismatches between acoustic realization and speakers’ intuitions for the same words collected in a previous study (Nadeu & Renwick 2016). As expected, results of the current and previous studies correlate well. Words that had mixed agreement showed larger numbers of mismatches.

4.2 Relevance to studies of bilingualism

The Catalan mid vowels have featured in a large number of studies on bilingualism. A common finding from these papers is that Catalan-dominant Catalan-Spanish bilinguals outperform Spanish-dominant bilinguals when perceiving the contrast (Bosch et al., 2000; Pallier et al., 1997, 2001). Our findings are relevant in light of this previous body of literature. We have shown that self-reported confidence in one's production of and distinction between the mid vowels correlates positively with dominance in Catalan. We have also demonstrated that these aspects correlate with strength of intuitions (gauged by certainty in height choice). It seems likely that increased sensitivity to the phonetic contrast results in greater ease in judging mid vowel height in individual words. Yet, it is also possible that stronger intuitions and ease of retrieval impact perception of the contrast.

In any case, the results presented here highlight the importance of taking into account individual speaker intuitions about their mental representations in certain phonetic studies. In fact, recent studies show that participants' intuitions correlate with their productions (Tilsen & Cohn, 2016). Bosch & Ramon-Casas (2011) examined front mid vowel production by two groups of speakers who differed in their L1. The initial analysis of formant data indicated that the L1 Spanish group produced a much smaller acoustic difference between the high and low mid vowels than the L1 Catalan group. Yet, when the same data were analyzed taking into consideration the actual vowel produced, not the one expected by the experimenters, no difference was found between the two groups in the phonetic implementation of the contrast. Thus, the authors show that, when accounting for differences in mental representations, the actual realization of the contrast does not vary across groups.

Bosch & Ramon-Casas (2011, p. 521) attribute the initial significant difference to "category selection errors", probably due to "unstable representations of these vowels in their lexicon". In discussing these and other findings, Simonet (2014, p. 36) talks about phonolexical representations being less robust or not accurate in Spanish-dominant speakers. The former terminology may be more appropriate, because it is unclear whether we can talk of "errors" in phonological category selection when there is so much interspeaker variability for some words.

The results we report here and in a previous study (Nadeu & Renwick, 2016) confirm the presence of variation in the lexical distribution of mid vowels across Catalan-dominant speakers of the same dialectal variety. Different lexical distribution of the mid vowels in some words across speakers might point to the fact that the mid vowel phonological contrasts are somewhat special, not as strong as other contrasts.

This interpretation is not new. A perception study (Mora et al., 2011) comparing the identification of low mid vs. high mid vowels to that of high mid vs. high vowels revealed that the former are perceived less categorically than the latter. A compressed S-shaped identification curve for the mid vowels was also found in Pallier et al. (1997), although that study did not compare the perception of mid vowels to that of other vowel pairs. Based on these and other findings, Mora, Keidel and Flege (2015) put forward the perceptual weakening hypothesis, which claims that the mid vowel contrasts are debilitated due to both internal and external factors.

4.3 Relevance to the notion of phonological contrast

As summarized in section 4.1, our study has confirmed accounts describing the mid vowels as highly conditioned by phonological context: native speaker intuitions favor open mid vowels in particular in a variety of contexts. What are the implications of this pattern for the mid vowel contrasts themselves? On the one hand, the effects of phonological conditioning help to maintain a distinction between the close and open vowels. Speakers remain highly aware of the presence of high mid and low mid qualities, although it is unlikely that they are aware of the contexts in which they select each one, for example that /ɛ, ɔ/ are preferred in syllables preceding unstressed [i]. However, it is important to note that when a particular sound is phonologically conditioned, its distribution becomes increasingly *predictable*, meaning that speakers need not depend on a lexical specification of, for instance, vowel height to produce or expect the appropriate phonetic form. In this situation, the presence of phonological conditioning effectively reduces the functional load of the mid vowel contrasts: because the close and open qualities are already restricted to particular contexts, there are few minimal pairs, and thus few lexical distinctions would be lost in the case of merger (Hockett, 1966; Surendran & Niyogi, 2006). Such a case could be formally treated by Hall's (2009) Probabilistic Phonological Relationship Model, an information-theoretic account of contrast based on predictability of distribution. As shown by Wedel et al. (2013), contrasts with lower functional load are more likely to be lost to merger over time. However as long as speaker awareness of the distinction persists, it is likely that among Catalan-dominant speakers, a phonetic distinction will also be maintained.

Another model exists for considering the effects of phonological conditioning on individual contrasts: the Multidimensional Model of Phonemic Robustness (MMPR) (Renwick, 2014). Proposed in light of the marginal contrast between Romanian /ɨ, ʌ/, this model incorporates three sets of factors that can contribute positively or negatively to "the degree to which a given phone is independent from other forces at work in the linguistic system" (Renwick, 2014, p. 182). A highly independent, robust phoneme is used frequently, with few restrictions, across a variety of phonological and morphological contexts, while a less independent phone appears in restricted contexts, to the extent that native speakers learn and rely on its co-occurrence with other sounds or structures. Phonological conditioning is one of three systemic factors identified by the MMPR, and it is argued to detract from a phone's independence, while the presence of lexical contrast and high type frequency would both have the opposite effect. Thus phonological conditioning and the scarcity of minimal pairs may detract from the independence of Catalan mid vowels from their typical contexts, while their type frequency across words (though not quantified here[11]) is likely high. It is also possible to evaluate the Catalan mid vowels with respect to the MMPR's phonetic factors: perceptibility, and phonetic distinctness. It is clear that among Catalan-dominant speakers, the close and open mid vowels remain highly perceptually distinct, contributing to their robustness. However, an acoustic study of the mid vowels' distribution (Nadeu & Renwick, 2016) has shown that for some speakers, the open and close vowels do not occupy distinct, non-overlapping portions of the vowel space; rather, they produce a single cloud that spans a wide range of F1 values. While tokens at the highest and lowest ranges are quite phonetically distinct, it can be unclear where the speaker's intended dividing line is between open and

close variants; thus the role of phonetic distinctness in maintaining the mid vowel contrasts is less clear. The MMPR also incorporates usage-based factors including morphologized uses (such as the heavy use of a phone in particular morphemes), which detract from robustness by increasing context-based predictability, and token frequency, which adds to a sound's robustness. In Catalan, the mid vowels are used in several frequent, stress-bearing morphological suffixes: /ɛ/ is found in the productive suffix *–et/eta* (diminutive, e.g. *caseta* 'little house'), while *–dor* (*pescador*, 'fisherman') has /o/, and *–ció* (*publicació*, 'publication') has /o/. Token frequency, like type frequency, has not to our knowledge been evaluated for Catalan, though we hypothesize that the mid vowels enjoy frequent usage.

The MMPR offers a holistic gauge of the strength of a sound category, and based on how – or in what way – a particular sound or contrast is robust, it is possible to make predictions. Specifically, if two sounds are not in a strong relationship of contrast, their phonetic realization and perception may be affected, particularly in contexts where one sound is highly predictable. In Romanian, acoustic analysis of continuous speech has shown that the marginally-contrastive /i̯, ʌ/ distinction is severely diminished, with a phonetic near-merger; the contrast indeed has a very low functional load, and results from automatic speech recognition experiments imply that a merger would have little detrimental impact on word recognition (Renwick, Vasilescu, Dutrey, Lamel, & Vieru, 2016).

In addition to filling out the portrait of how the Catalan mid vowel contrasts operate, our findings build on previous results that combined speaker-specific phonological intuitions with individual vowel production (Nadeu & Renwick, 2016). In that paper, we argued that, across 100 words with stressed mid vowels, there was only moderate inter-speaker agreement on vowel height; and that in a significant number of cases, speakers' intuitions did not match the acoustics of their productions. The present study confirms, however, that there are words where there is little variability in judgments; these words, which nearly always contain open mid vowels, are very stable. In the three Nonvariable words that were included in both studies, it is clear that speakers had few intuition-production mismatches. These were: *aterra* '(it) lands', which was judged to have /ɛ/ by 12/14 speakers (with 6/41 acoustics-judgment mismatches); *perdre* 'to lose' (12/14 /ɛ/ judgments, with 6/41 acoustics-judgment mismatches), and *verda* 'green, f.' (13/14 /ɛ/ judgments, with 3/41 acoustics-judgment mismatches). These results lead us to hypothesize that the participants in the present survey would also produce these words accurately (according to their intuition), with open vowels. It also seems to be the case that among words without a specific context conditioning mid vowel height, some words are more variable than others. We do not have a full explanation of why these words are so variable, and more research is needed upon this point.

It seems that the contrasts among mid vowels, with respect to other vowels in the Catalan system, are less robust. It is easy to argue that this is a change due to increased contact with Spanish, but in fact, similar patterns emerge across the Romance languages. As discussed in Nadeu and Renwick (2016), certain varieties of French, Italian, and Portuguese, for instance, are undergoing weakening of the mid vowel contrasts, meaning that they are partially neutralized or variable across speakers (Dufour, Nguyen, & Frauenfelder, 2007; Fagyal, Hassa, & Ngom, 2002; Renwick & Ladd, 2016; Wetzels,

1992). Similarly, not all Catalan varieties maintain both front and back mid vowel contrasts (Recasens & Espinosa, 2009).

5. Conclusion

This study has confirmed that there is some phonological conditioning of mid vowels in Catalan, in line with previous descriptions of vowel quality in the language. Where phonological conditioning occurs, speakers judge quality with increased consistency and confidence. It remains an open question, which we leave for future research, whether this phonological conditioning is synchronically active or restricted to a closed class of items for historical reasons. Other words that lack phonological conditioning environments, however, are more variable in mid vowel height. A second open research question is *why* some words are more variable than others; one potential factor is lexical frequency, as well as the source of the word (e.g. neologisms vs. native Romance words; Cabré, 2009), or the effects of Spanish cognates on speakers' phonolexical representations. We have shown that vowel choice is related to speaker age, and certainty of judgment is related to language dominance. Specifically, older participants select high mid vowels more frequently. On the other hand, participants with higher Catalan dominance have greater confidence in their judgments, selecting "Clearly Close" or "Clearly Open" more often than participants with lower BLP scores. In general, though, a high degree of certainty is found regardless of Catalan dominance level.

Our results suggest that the robustness of phonolexical representations may depend on language dominance but also on word identity. Despite high levels of agreement on many words, especially those with phonological conditioning, our survey confirms that variation in mid vowel height is present in Catalan, to the point that even naïve speakers are aware of it. Since this variation has reached the threshold of speaker awareness, we argue that it affects the vowels' encoding in phonological representations, for instance broadening the phonetic range associated with their exemplar clouds. We also argue that in cases where vowel quality is highly predictable, for instance in our Nonvariable stimuli, the vowel specification in speakers' representations may be less important simply because context dictates vowel height. This result in particular builds on previous findings that some words are variably produced even by the same speaker, and those words elicit greater variation in judgments.

This research contributes to the larger question of how phonology and phonetics map to one another. Phonological contrasts are traditionally treated as imposing discrete distinctions on a phonetic continuum, but our results demonstrate that apart from phonetic continuity, the phonological distinction itself is not robust in all contexts. The continuous nature of vowel acoustics contributes to the complexity of studying the relationship between their production and speakers' intuitions, which is a common theme across the studies mentioned here. When productions are collected, they must still be classified, which may be done by the researcher either aurally or statistically. However because there can be substantial acoustic overlap of vowel distributions, it is possible that statistical methods introduce misclassifications; likewise, a listener's own language experience can induce perceptual bias to an aural classification method. An alternative, which our research program explores, is to rely on speakers' judgments, which has the

benefit of accessing their phonological intuitions. By using multiple techniques to classify vowels (e.g. speaker intuitions *and* statistical or aural confirmation), we shed new light on the nature of Catalan mid vowel contrasts specifically, and offer methods for increased precision in mapping abstract representations to acoustic reality.

Acknowledgements
The authors are grateful for the input of Miquel Simonet, the Editors and Associate Editor, and one anonymous reviewer; and for the participation of the Catalan speakers and colleagues who completed and distributed the survey itself.

Funding
This article was supported by funds from the University of Georgia Franklin College of Arts and Sciences, and from the Department of Spanish, Italian and Portuguese at Penn State, to which the second author was affiliated.

Notes

1. Catalonia is a bilingual community. Catalan speakers are bilingual in Spanish, although the time at which each language is learned and the degree of dominance in each language varies across speakers. The 2013 language census report (Institut d'Estadística de Catalunya, 2015) shows that 94.3% of the population understands Catalan, while 99.8% understand Spanish. As far as speaking the language is concerned, 99.7% of the population self-reported being able to speak Spanish, whereas only 80.4% claim that they can speak Catalan.

2. The survey and its results can be accessed at http://www.phon.ucl.ac.uk/home/wells/poll98.htm.

3. Prestressing suffixes are unstressed suffixes that move the position of the stress to the last syllable in the root (*símbol* [ˈsimbul] 'symbol', but *simbòlic* [simˈbɔlik] 'symbolic').

4. We use "NA" to indicate that the Immediate Context is not relevant for back mid vowels.

5. The use of "0" reflects the fact that it is difficult to find front mid vowels that are conditioned *only* by a prestressing suffix, and not also by posttonic [i, j].

6. Shortly after our study was conducted in winter 2016, the Institut d'Estudis Catalans in fact revised Catalan spelling conventions to exclude most diacritic stress marks on minimal pairs (Institut d'Estudis Catalans, 2016). Our stimuli included three minimal pairs, some of which formerly required a diacritic mark, but no longer do.

7. The range of completion times is so broad because participants were allowed to complete the survey in multiple sessions. That is, they could stop their participation at any time and resume it at their convenience by returning to the Qualtrics website. We observe a few cases of clear outliers (> 500 minutes). The median duration was 21.7 minutes.

8. By contrast, running each condition separately produced uninterpretable results due to the *lack* of variation in responses to Nonvariable words; for an explanation see Sim & Wright (2005).

9. The results are different if Spanish-dominant participants are included: approximately 30% more words, across the three conditions, have at least one ambiguous response: in the Nonvariable condition, only 25 words have no ambiguous responses; No Prediction, only 10 words; and Variable, only 11 words have no ambiguous responses.

10. A z-score transformation results in values centered on a mean of 0, indicating the number of standard deviations that separate the raw value from the mean. For instance, the mean BLP score 69.06 becomes 0, and a score that is one standard deviation away, i.e. 47.74 points lower or higher, has a scaled value of either –1 or +1, respectively.

11. Rafel i Fontanals (1980) presents frequency data on the phonemes of Catalan, based on counts from a 10,000-word literary text. He finds that the mid vowels are less frequent than stressed /i/ and /a/, but more frequent than stressed /u/, so they are not rare. However, as discussed elsewhere (Nadeu & Renwick, 2016, p. 39), the presence of interspeaker variability in vowel height judgment and production complicates the calculation of accurate frequency measures.

Appendix A. Stimuli

List 1 Stimuli

Number	Word	Translation	Backness	Condition
1	accessori	accessory	Back	Nonvariable
2	al·lucinogen	hallucinogenic	Back	Nonvariable
3	bròquil	broccoli	Back	Nonvariable
4	cardiòleg	cardiologist	Back	Nonvariable
5	cefalòpode	cephalopod	Back	Nonvariable
6	cònic	conical	Back	Nonvariable
7	coreògraf	choreographer	Back	Nonvariable
8	dimoni	devil	Back	Nonvariable
9	divorci	divorce	Back	Nonvariable
10	egòlatra	narcissist	Back	Nonvariable
11	història	story	Back	Nonvariable
12	introductori	introductory	Back	Nonvariable
13	lacrimogen	tear-producing	Back	Nonvariable
14	metòdic	methodical	Back	Nonvariable
15	odi	hatred	Back	Nonvariable
16	ossi	osseous	Back	Nonvariable
17	propi	own	Back	Nonvariable
18	robatori	theft	Back	Nonvariable
19	sociòleg	sociologist	Back	Nonvariable
20	territori	territory	Back	Nonvariable
21	compota	compote	Back	Tested (No Prediction)
22	espora	spore	Back	Tested (No Prediction)

Number	Word	Translation	Backness	Condition
23	fora	outside	Back	Tested (No Prediction)
24	fotre	do, pop.	Back	Tested (No Prediction)
25	postres	dessert	Back	Tested (No Prediction)
26	proposa	proposes	Back	Tested (No Prediction)
27	prosa	prose	Back	Tested (No Prediction)
28	rossa	blonde, f.	Back	Tested (No Prediction)
29	sostre	ceiling	Back	Tested (Variable)
30	torça	bends	Back	Tested (No Prediction)
31	brossa	trash	Back	Tested (No Prediction)
32	crosta	scab/crust	Back	Tested (Variable)
33	dona	woman	Back	Tested (No Prediction)
34	dona	gives	Back	Tested (No Prediction)
35	esborra	deletes/erases	Back	Tested (Variable)
36	flora	flora	Back	Tested (No Prediction)
37	fossa	grave, pit	Back	Tested (No Prediction)
38	lactosa	lactose	Back	Tested (No Prediction)
39	lectora	reader, f.	Back	Tested (No Prediction)
40	polpa	pulp	Back	Tested (No Prediction)
41	barbacoa	bbq	Back	Variable
42	broma	joke	Back	Variable
43	cofoi	proud	Back	Variable
44	decora	decorates	Back	Variable
45	dojo	abundantly	Back	Variable
46	dropo	lazy person	Back	Variable
47	estona	(a) while	Back	Variable
48	flonjo	soft	Back	Variable
49	gol	goal	Back	Variable
50	hora	hour	Back	Variable
51	incomoda	makes uncomfortable	Back	Variable
52	mormola	whispers	Back	Variable
53	mot	word	Back	Variable
54	norma	norm	Back	Variable
55	penjoll	pendant	Back	Variable
56	proa	prow	Back	Variable
57	rebolca	rolls around	Back	Variable
58	rostre	face	Back	Variable
59	sorra	sand	Back	Variable
60	vota	votes	Back	Variable
61	alfabètic	alphabetic	Front	Nonvariable
62	ciència	science	Front	Nonvariable
63	espècia	spice	Front	Nonvariable
64	èxit	success	Front	Nonvariable
65	gel	ice	Front	Nonvariable
66	gerro	vase	Front	Nonvariable
67	guerra	war	Front	Nonvariable
68	higiènic	hygienic	Front	Nonvariable
69	imbècil	idiot	Front	Nonvariable
70	lleig	ugly	Front	Nonvariable
71	llenya	firewood	Front	Nonvariable
72	magisteri	teaching education	Front	Nonvariable
73	mel	honey	Front	Nonvariable
74	molèstia	annoyance, discomfort	Front	Nonvariable
75	orenga	oregano	Front	Nonvariable
76	perla	pearl	Front	Nonvariable
77	renya	scolds	Front	Nonvariable

Number	Word	Translation	Backness	Condition
78	tebi	lukewarm	Front	Nonvariable
79	terra	land/earth	Front	Nonvariable
80	violència	violence	Front	Nonvariable
81	aterra	lands	Front	Tested (Nonvariable)
82	cercle	circle	Front	Tested (No Prediction)
83	esfera	sphere	Front	Tested (Variable)
84	peça	piece	Front	Tested (No Prediction)
85	perdre	to lose	Front	Tested (Nonvariable)
86	pesa.n	weight	Front	Tested (Variable)
87	pesa.v	weighs (v)	Front	Tested (Variable)
88	pesta	plague	Front	Tested (No Prediction)
89	setze	sixteen	Front	Tested (No Prediction)
90	vespa	wasp	Front	Tested (No Prediction)
91	barrera	barrier	Front	Tested (No Prediction)
92	cera	wax	Front	Tested (No Prediction)
93	enceta	begins (specific)	Front	Tested (Variable)
94	festa	party	Front	Tested (No Prediction)
95	neta	clean	Front	Tested (No Prediction)
96	neta	granddaughter	Front	Tested (No Prediction)
97	pera	pear	Front	Tested (No Prediction)
98	Pere	Peter	Front	Tested (No Prediction)
99	sencera	entire	Front	Tested (No Prediction)
100	verda	green	Front	Tested (Nonvariable)
101	anem	(we) go	Front	Variable
102	arregla	fixes	Front	Variable
103	comença	starts	Front	Variable
104	consell	advice	Front	Variable
105	decebre	disappoint	Front	Variable
106	dialecte	dialect	Front	Variable
107	diumenge	Sunday	Front	Variable
108	entendre	understand	Front	Variable
109	escena	scene	Front	Variable
110	esquerp	fierce	Front	Variable
111	feble	weak	Front	Variable
112	hotel	hotel	Front	Variable
113	interessa	interests	Front	Variable
114	llengua	language/tongue	Front	Nonvariable
115	membre	member	Front	Variable
116	penja	hangs	Front	Variable
117	primavera	Spring	Front	Variable
118	renta	washes	Front	Variable
119	rovell	yolk	Front	Variable
120	sistema	system	Front	Variable

List 2 Stimuli

Number	Word	Translation	Backness	Condition
121	aeròdrom	airdrome	Back	Nonvariable
122	bigoti	moustache	Back	Nonvariable
123	canòdrom	dog track	Back	Nonvariable
124	carquinyoli	type of cookie	Back	Nonvariable
125	codi	code	Back	Nonvariable
126	consultori	doctor's office	Back	Nonvariable

Number	Word	Translation	Backness	Condition
127	diagnosi	diagnose	Back	Nonvariable
128	divisori	dividing	Back	Nonvariable
129	dosi	dose	Back	Nonvariable
130	geògraf	geographer	Back	Nonvariable
131	indòmit	indomitable	Back	Nonvariable
132	ivori	ivory	Back	Nonvariable
133	matrimoni	marriage	Back	Nonvariable
134	negoci	business	Back	Nonvariable
135	oli	oil	Back	Nonvariable
136	podòleg	podiatrist	Back	Nonvariable
137	psicòleg	psychologist	Back	Nonvariable
138	soci	member	Back	Nonvariable
139	somni	dream	Back	Nonvariable
140	xenòfob	xenophobic	Back	Nonvariable
141	cotorra	parrot	Back	Tested (No Prediction)
142	estora	carpeta	Back	Tested (No Prediction)
143	forta	strong, f.	Back	Tested (No Prediction)
144	grossa	large, f.	Back	Tested (No Prediction)
145	orca	orca	Back	Tested (No Prediction)
146	pota	leg (animal)	Back	Tested (No Prediction)
147	rebota	bounces	Back	Tested (No Prediction)
148	reposa	rests	Back	Tested (No Prediction)
149	sorda	deaf, f.	Back	Tested (No Prediction)
150	torra	toasts	Back	Tested (No Prediction)
151	brossa	trash	Back	Tested (No Prediction)
152	crosta	scab/crust	Back	Tested (Variable)
153	dona	woman	Back	Tested (No Prediction)
154	dona	gives	Back	Tested (No Prediction)
155	esborra	deletes/erases	Back	Tested (Variable)
156	flora	flora	Back	Tested (No Prediction)
157	fossa	grave, pit	Back	Tested (No Prediction)
158	lactosa	lactose	Back	Tested (No Prediction)
159	lectora	reader, f.	Back	Tested (No Prediction)
160	polpa	pulp	Back	Tested (No Prediction)
161	boa	boa	Back	Variable
162	brotxa	brush	Back	Variable
163	col·loca	puts	Back	Variable
164	destorba	bothers	Back	Variable
165	dota	provides	Back	Variable
166	esbossa	sketches	Back	Variable
167	estrofa	stanza	Back	Variable
168	flor	flower	Back	Variable
169	hola	hello	Back	Variable
170	idioma	language	Back	Variable
171	molsa	moss	Back	Variable
172	mos	bite (n)	Back	Variable
173	noces	wedding	Back	Variable
174	olla	pot	Back	Variable
175	por	fear	Back	Variable
176	provoca	provokes	Back	Variable
177	rodola	falls spinning	Back	Variable
178	sobra	is not necessary	Back	Variable
179	udol	howls	Back	Variable
180	zona	zone	Back	Variable
181	cel	sky	Front	Nonvariable

Number	Word	Translation	Backness	Condition
182	encanteri	spell	Front	Nonvariable
183	estel	star	Front	Nonvariable
184	expert	expert	Front	Nonvariable
185	geni	genius	Front	Nonvariable
186	gresca	party	Front	Nonvariable
187	herència	inheritance	Front	Nonvariable
188	homogeni	homogeneous	Front	Nonvariable
189	juguesca	bet	Front	Nonvariable
190	lleixa	shelf	Front	Nonvariable
191	llengua	tongue, language	Front	Nonvariable
192	medi	environment	Front	Nonvariable
193	misteri	mystery	Front	Nonvariable
194	obert	open	Front	Nonvariable
195	paciència	patience	Front	Nonvariable
196	règim	diet	Front	Nonvariable
197	silenci	silence	Front	Nonvariable
198	tel	thin layer	Front	Nonvariable
199	vel	veil	Front	Nonvariable
200	nervi	nerve	Front	Nonvariable
201	carpeta	folder	Front	Tested (No Prediction)
202	cresta	comb (birds)	Front	Tested (Variable)
203	espessa	thick	Front	Tested (No Prediction)
204	pantera	panther	Front	Tested (No Prediction)
205	pedra	stone	Front	Tested (No Prediction)
206	pesca.v	fishes	Front	Tested (Nonvariable)
207	pesca.n	fishing	Front	Tested (Nonvariable)
208	pressa	hurry	Front	Tested (No Prediction)
209	tretze	thirteen	Front	Tested (No Prediction)
210	zeta	zee	Front	Tested (No Prediction)
211	barrera	barrier	Front	Tested (No Prediction)
212	cera	wax	Front	Tested (No Prediction)
213	enceta	begins (specific)	Front	Tested (Variable)
214	festa	party	Front	Tested (No Prediction)
215	neta	clean	Front	Tested (No Prediction)
216	neta	granddaughter	Front	Tested (No Prediction)
217	pera	pear	Front	Tested (No Prediction)
218	Pere	Peter	Front	Tested (No Prediction)
219	sencera	entire	Front	Tested (No Prediction)
220	verda	green	Front	Tested (Nonvariable)
221	aquest	this	Front	Variable
222	aspecte	aspect	Front	Variable
223	concret	specific	Front	Variable
224	crea	creates	Front	Variable
225	desa	saves/puts away	Front	Variable
226	dilema	dilemma	Front	Variable
227	ell	he	Front	Variable
228	entre	between	Front	Variable
229	esquema	scheme/outline	Front	Variable
230	exemple	example	Front	Variable
231	fem	(we) do	Front	Variable
232	insecte	insect	Front	Variable
233	llei	law	Front	Variable
234	marea	tide	Front	Variable
235	mes	month	Front	Variable
236	perfecte	perfect	Front	Variable

Number	Word	Translation	Backness	Condition
237	problema	problem	Front	Variable
238	romesco	type of sauce	Front	Variable
239	sense	without	Front	Variable
240	tren	train	Front	Variable

Appendix B. Questions about speakers' intuitions

1. Creus que pronuncies la diferència entre e oberta i e tancada? Indica el teu grau de certesa.
   "Do you think you pronounce the difference between open "e" and close "e"? Indicate your degree of certainty."
2. Creus que pronuncies la diferència entre o oberta i o tancada? Indica el teu grau de certesa.
   "Do you think you pronounce the difference between open "o" and close "o"? Indicate your degree of certainty."
3. Quan penses en paraules que tenen una *e* o una *o,* t'és fàcil saber si la vocal és oberta o tancada? Indica el grau de facilitat.
   "When you think about words that include an "e" or an "o", is it easy for you to judge whether the vowel is open or close? Indicate your degree of certainty."
4. T'has adonat mai que altres persones poden utilitzar aquestes vocals de manera diferent de com les utilitzes tu? Per exemple, t'has adonat mai que algú pronuncia una paraula amb e tancada i tu ho fas amb e oberta?
   "Have you ever noticed that other people may use these vowels in a different manner from how you use them? For example, have you ever noticed that someone pronounces a word with close "e" while you pronounce it with open "e"?

Appendix C. Summary of participant judgments

Word	% Close judgments	% Open judgments	% Other judgments	Condition
accessori	10.00	90.00	0.00	Nonvariable
aeròdrom	1.47	95.59	2.94	Nonvariable
al·lucinogen	5.00	90.00	5.00	Nonvariable
alfabètic	0.00	100.00	0.00	Nonvariable
anem	63.33	28.33	8.33	Variable
aquest	39.71	60.29	0.00	Variable
arregla	93.33	6.67	0.00	Variable
aspecte	82.35	16.18	1.47	Variable
aterra	6.67	90.00	3.33	Nonvariable
barbacoa	90.00	8.33	1.67	Variable
barrera	90.62	7.03	2.34	No prediction
bigoti	7.35	89.71	2.94	Nonvariable
boa	94.12	5.88	0.00	Variable
broma	93.33	5.00	1.67	Variable
bròquil	5.00	93.33	1.67	Nonvariable
brossa	8.59	90.62	0.78	No prediction

Word	% Close judgments	% Open judgments	% Other judgments	Condition
brotxa	27.94	67.65	4.41	Variable
canòdrom	5.88	92.65	1.47	Nonvariable
cardiòleg	5.00	95.00	0.00	Nonvariable
carpeta	4.41	92.65	2.94	No prediction
carquinyoli	5.88	91.18	2.94	Nonvariable
cefalòpode	8.33	86.67	5.00	Nonvariable
cel	1.47	97.06	1.47	Nonvariable
cera	7.81	92.19	0.00	No prediction
cercle	95.00	5.00	0.00	No prediction
ciència	5.00	95.00	0.00	Nonvariable
codi	4.41	92.65	2.94	Nonvariable
cofoi	18.33	76.67	5.00	Variable
col·loca	80.88	14.71	4.41	Variable
comença	38.33	58.33	3.33	Variable
compota	11.67	86.67	1.67	No prediction
concret	91.18	5.88	2.94	Variable
cònic	5.00	93.33	1.67	Nonvariable
consell	41.67	46.67	11.67	Variable
consultori	8.82	89.71	1.47	Nonvariable
cotorra	89.71	8.82	1.47	No prediction
crea	85.29	13.24	1.47	Variable
cresta	27.94	70.59	1.47	Variable
crosta	17.19	78.91	3.91	Variable
decebre	16.67	83.33	0.00	Variable
decora	86.67	10.00	3.33	Variable
desa	57.35	35.29	7.35	Variable
destorba	92.65	2.94	4.41	Variable
diagnosi	1.47	97.06	1.47	Nonvariable
dialecte	6.67	93.33	0.00	Variable
dilema	11.76	83.82	4.41	Variable
dimoni	8.33	88.33	3.33	Nonvariable
diumenge	95.00	3.33	1.67	Variable
divisori	13.24	85.29	1.47	Nonvariable
divorci	6.67	86.67	6.67	Nonvariable
dojo	51.67	40.00	8.33	Variable
dona (n)	7.81	90.62	1.56	No prediction
dona (v)	91.41	8.59	0.00	No prediction
dosi	2.94	95.59	1.47	Nonvariable
dota	86.76	5.88	7.35	Variable
dropo	56.67	38.33	5.00	Variable
egòlatra	10.00	86.67	3.33	Nonvariable
ell	88.24	8.82	2.94	Variable
encanteri	1.47	97.06	1.47	Nonvariable
enceta	50.00	46.88	3.12	Variable
entendre	75.00	21.67	3.33	Variable
entre	79.41	19.12	1.47	Variable
esborra	94.53	4.69	0.78	Variable
esbossa	88.24	7.35	4.41	Variable
escena	6.67	93.33	0.00	Variable
esfera	88.33	11.67	0.00	Variable
espècia	6.67	90.00	3.33	Nonvariable
espessa	10.29	89.71	0.00	No prediction
espora	88.33	11.67	0.00	No prediction
esquema	83.82	14.71	1.47	Variable
esquerp	20.00	76.67	3.33	Variable

Word	% Close judgments	% Open judgments	% Other judgments	Condition
estel	5.88	94.12	0.00	Nonvariable
estona	90.00	8.33	1.67	Variable
estora	92.65	4.41	2.94	No prediction
estrofa	7.35	89.71	2.94	Variable
exemple	94.12	5.88	0.00	Variable
èxit	1.67	98.33	0.00	Nonvariable
expert	14.71	85.29	0.00	Nonvariable
feble	93.33	5.00	1.67	Variable
fem	42.65	50.00	7.35	Variable
festa	92.19	7.03	0.78	No prediction
flonjo	45.00	48.33	6.67	Variable
flor	10.29	82.35	7.35	Variable
flora	49.22	42.97	7.81	No prediction
fora	15.00	83.33	1.67	No prediction
forta	10.29	88.24	1.47	No prediction
fossa	95.31	3.12	1.56	No prediction
fotre	88.33	11.67	0.00	No prediction
gel	5.00	95.00	0.00	Nonvariable
geni	8.82	91.18	0.00	Nonvariable
geògraf	2.94	97.06	0.00	Nonvariable
gerro	1.67	96.67	1.67	Nonvariable
gol	10.00	85.00	5.00	Variable
gresca	1.47	97.06	1.47	Nonvariable
grossa	5.88	92.65	1.47	No prediction
guerra	10.00	90.00	0.00	Nonvariable
herència	4.41	91.18	4.41	Nonvariable
higiènic	1.67	96.67	1.67	Nonvariable
història	3.33	96.67	0.00	Nonvariable
hola	17.65	75.00	7.35	Variable
homogeni	4.41	95.59	0.00	Nonvariable
hora	11.67	86.67	1.67	Variable
hotel	28.33	70.00	1.67	Variable
idioma	79.41	19.12	1.47	Variable
imbècil	0.00	100.00	0.00	Nonvariable
incomoda	18.33	78.33	3.33	Variable
indòmit	1.47	85.29	13.24	Nonvariable
insecte	80.88	19.12	0.00	Variable
interessa	20.00	71.67	8.33	Variable
introductori	10.00	90.00	0.00	Nonvariable
ivori	2.94	91.18	5.88	Nonvariable
juguesca	1.47	98.53	0.00	Nonvariable
lacrimogen	10.00	86.67	3.33	Nonvariable
lactosa	91.41	7.81	0.78	No prediction
lectora	91.41	7.03	1.56	No prediction
llei	82.35	13.24	4.41	Variable
lleig	95.00	5.00	0.00	Nonvariable
lleixa	80.88	11.76	7.35	Nonvariable
llengua	66.41	30.47	3.12	Variable
llenya	95.00	5.00	0.00	Nonvariable
magisteri	1.67	98.33	0.00	Nonvariable
marea	83.82	11.76	4.41	Variable
matrimoni	2.94	92.65	4.41	Nonvariable
medi	7.35	92.65	0.00	Nonvariable
mel	8.33	90.00	1.67	Nonvariable
membre	96.67	3.33	0.00	Variable

Word	% Close judgments	% Open judgments	% Other judgments	Condition
mes	76.47	20.59	2.94	Variable
metòdic	1.67	96.67	1.67	Nonvariable
misteri	2.94	95.59	1.47	Nonvariable
molèstia	16.67	81.67	1.67	Nonvariable
molsa	91.18	7.35	1.47	Variable
mormola	40.00	40.00	20.00	Variable
mos	80.88	14.71	4.41	Variable
mot	90.00	8.33	1.67	Variable
negoci	2.94	97.06	0.00	Nonvariable
nervi	13.24	85.29	1.47	Nonvariable
neta (n)	5.47	93.75	0.78	No prediction
neta (adj)	83.59	15.62	0.78	No prediction
noces	17.65	80.88	1.47	Variable
norma	88.33	10.00	1.67	Variable
obert	10.29	88.24	1.47	Nonvariable
odi	3.33	96.67	0.00	Nonvariable
oli	0.00	97.06	2.94	Nonvariable
olla	16.18	82.35	1.47	Variable
orca	22.06	76.47	1.47	No prediction
orenga	3.33	95.00	1.67	Nonvariable
ossi	3.33	95.00	1.67	Nonvariable
paciència	0.00	97.06	2.94	Nonvariable
pantera	86.76	10.29	2.94	No prediction
peça	93.33	6.67	0.00	No prediction
pedra	92.65	4.41	2.94	No prediction
penja	93.33	6.67	0.00	Variable
penjoll	21.67	75.00	3.33	Variable
pera	5.47	93.75	0.78	No prediction
perdre	6.67	93.33	0.00	Nonvariable
Pere	91.41	7.03	1.56	No prediction
perfecte	82.35	16.18	1.47	Variable
perla	3.33	96.67	0.00	Nonvariable
pesa.n	18.33	81.67	0.00	Variable
pesa.v	60.00	38.33	1.67	Variable
pesca.n	5.88	91.18	2.94	Nonvariable
pesca.v	7.35	91.18	1.47	Nonvariable
pesta	88.33	10.00	1.67	No prediction
podòleg	0.00	100.00	0.00	Nonvariable
polpa	85.16	12.50	2.34	No prediction
por	95.59	2.94	1.47	Variable
postres	13.33	81.67	5.00	No prediction
pota	8.82	88.24	2.94	No prediction
pressa	85.29	14.71	0.00	No prediction
primavera	68.33	30.00	1.67	Variable
proa	90.00	10.00	0.00	Variable
problema	1.47	97.06	1.47	Variable
propi	8.33	88.33	3.33	Nonvariable
proposa	3.33	91.67	5.00	No prediction
prosa	6.67	88.33	5.00	No prediction
provoca	92.65	5.88	1.47	Variable
psicòleg	1.47	98.53	0.00	Nonvariable
rebolca	85.00	10.00	5.00	Variable
rebota	89.71	7.35	2.94	No prediction
règim	2.94	95.59	1.47	Nonvariable
renta	81.67	18.33	0.00	Variable

Word	% Close judgments	% Open judgments	% Other judgments	Condition
renya	15.00	83.33	1.67	Nonvariable
reposa	13.24	83.82	2.94	No prediction
robatori	15.00	83.33	1.67	Nonvariable
rodola	25.00	67.65	7.35	Variable
romesco	11.76	86.76	1.47	Variable
rossa	88.33	11.67	0.00	No prediction
rostre	46.67	50.00	3.33	Variable
rovell	61.67	38.33	0.00	Variable
sencera	17.97	78.91	3.12	No prediction
sense	69.12	30.88	0.00	Variable
setze	30.00	70.00	0.00	No prediction
silenci	5.88	91.18	2.94	Nonvariable
sistema	65.00	31.67	3.33	Variable
sobra	36.76	58.82	4.41	Variable
soci	7.35	89.71	2.94	Nonvariable
sociòleg	1.67	95.00	3.33	Nonvariable
somni	13.24	83.82	2.94	Nonvariable
sorda	95.59	2.94	1.47	No prediction
sorra	93.33	6.67	0.00	Variable
sostre	45.00	53.33	1.67	Variable
tebi	16.67	83.33	0.00	Nonvariable
tel	1.47	98.53	0.00	Nonvariable
terra	1.67	98.33	0.00	Nonvariable
territori	8.33	91.67	0.00	Nonvariable
torça	11.67	85.00	3.33	No prediction
torra	91.18	8.82	0.00	No prediction
tren	23.53	67.65	8.82	Variable
tretze	95.59	2.94	1.47	No prediction
udol	19.12	75.00	5.88	Variable
vel	8.82	89.71	1.47	Nonvariable
verda	4.69	95.31	0.00	Nonvariable
vespa	88.33	10.00	1.67	No prediction
violència	5.00	93.33	1.67	Nonvariable
vota	95.00	5.00	0.00	Variable
xenòfob	1.47	95.59	2.94	Nonvariable
zeta	7.35	92.65	0.00	No prediction
zona	86.76	7.35	5.88	Variable

References

Alkire, T., & Rosen, C. (2010). *Romance Languages: A Historical Introduction*. Cambridge: Cambridge University Press.

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.

Bates, D., & Maechler, M. (2009). lme4: Linear mixed-effects models using S4 classes (Version R package version 0.999375-32).

Berko, J. (1958). The Child's Learning of English Morphology. *WORD*, *14*(2–3), 150–177. https://doi.org/10.1080/00437956.1958.11659661

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*. Retrieved from https://sites.la.utexas.edu/bilingual/

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.

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.

Carrera-Sabaté, J. (2005). Les vocals àtones medials dels compostos en lleidatà, entre la variació i la imperceptibilitat. *Estudis de Llengua i Literatura Catalanes*, *7*, 251–272.

Carrera-Sabaté, J. (2006). Some connections between linguistic change and the written language: The behavior of speakers aged 3 to 20. *Language Variation and Change*, *18*(01), 15–34. https://doi.org/10.1017/S0954394506060017

Cassidy, F., & Hall, J. (1985). *The Dictionary of American Regional English, Volumes I-IV*. Cambridge, MA: Harvard University Press.

Chambers, J. K. (1994). An introduction to dialect topography. *English World-Wide*, *15*(1), 35–53.

Chambers, J. K. (1998a). Inferring dialect from a postal questionnaire. *Journal of English Linguistics*, *26*(3), 222–246.

Chambers, J. K. (1998b). Social embedding of changes in progress. *Journal of English Linguistics*, *26*(1), 5–36.

Coromines, J. (1974). *Lleures i converses d'un filòleg*. Barcelona: Club Editor.

Côté, M.-H., & Kharlamov, V. (2010). The impact of experimental tasks on syllabification judgments: A case study of Russian. In C. E. Cairns & E. Raimy (Eds.), *Handbook of the Syllable* (pp. 271–294). Leiden: Brill.

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.

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: Cambridge University Press.

Fleiss, J. L. (1971). Measuring nominal scale agreement among many raters. *Psychological Bulletin*, *76*(5), 378–382.

Frisch, S. A., Large, N. R., & Pisoni, D. B. (2000). Perception of wordlikeness: Effects of segment probability and length on the processing of non-words. *Journal of Memory and Language*, *42*(4), 481–496.

Frisch, S. A., & Zawaydeh, B. A. (2001). The psychological reality of OCP-Place in Arabic. *Language*, *77*(1), 91–106.

García, C. M. (2015). *Gradience and Variability of Intervocalic /s/ Voicing in Highland Ecuadorian Spanish* (Ph.D. Dissertation). The Ohio State University, Columbus, OH.

Hall, K. C. (2009). *A Probabilistic Model of Phonological Relationships from Contrast to Allophony* (Ph.D. Dissertation). The Ohio State University, Columbus, OH.

Hay, J., Pierrehumbert, J. B., & Beckman, M. E. (2003). Speech perception, well-formedness and the statistics of the lexicon. In J. Local, R. Ogden, & R. Temple (Eds.), *Papers in Laboratory Phonology VI: Phonetic Interpretation* (pp. 58–74). Cambridge: Cambridge University Press.

Henriksen, N. (2017). Patterns of vowel laxing and harmony in Iberian Spanish: Data from production and perception. *Journal of Phonetics*, *63*(Supplement C), 106–126. https://doi.org/10.1016/j.wocn.2017.05.001

Hockett, C. F. (1966). *The quantification of functional load: A linguistic problem* (Memorandum No. RM-5168-PR).

Institut d'Estadística de Catalunya. (2015). *Enquesta d'usos lingüístics de la població. 2013 [Survey of the population's language uses. 2013]*. Barcelona: Generalitat de Catalunya. Retrieved from http://idescat.cat/p/eulp2013

Institut d'Estudis Catalans. (2016). *Gramàtica de la llengua catalana*. Barcelona, Spain: Institut d'Estudis Catalans.

Kurath, H., & McDavid, R. I. J. (1961). *The Pronunciation of English in the Atlantic States: Based upon the Collections of the Linguistic Atlas of the Eastern United States*. University of Michigan Press.

Labov, W., Ash, S., & Boberg, C. (2006). *The Atlas of North American English: Phonetics, phonology, and sound change: A multimedia reference tool*. Berlin; New York: Mouton de Gruyter.

Labov, W., Karen, M., & Miller, C. (1991). Near-mergers and the suspension of phonemic contrast. *Language Variation and Change*, *3*(1), 33–74. https://doi.org/10.1017/S0954394500000442

Labov, W., Yaeger, M., & Steiner, R. (1972). *A quantitative study of sound change in progress* (Vol. 1). US Regional Survey.

Landick, M. (1995). The mid-vowels in figures: hard facts. *French Review*, *69*(1), 88–102.

Lleó, C., Benet, A., & Cortés, S. (2007). Some current phonological features in the Catalan of Barcelona. *Catalan Review*, *21*(1), 279–300.

Martinet, A. (1945). *La prononciation du français contemporain: Témoignages recueillis en 1941 dans un camp d'officiers prisonniers*. Paris: Droz.

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. 0079–102).

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.

McDavid, R. I. J., & O'Cain, R. K. (1980). *Linguistic Atlas of the Middle and South Atlantic States* (Vol. 1). Chicago: University of Chicago Press.

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., & Renwick, M. E. L. (2016). Variation in the lexical distribution and implementation of phonetically similar phonemes in Catalan. *Journal of Phonetics*, *58*, 22–47. https://doi.org/10.1016/j.wocn.2016.05.003

Nagy, N. (2001). Live Free or Die as a linguistic principle. *American Speech*, *76*(1), 30–41.

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.
Pederson, L., McDaniel, S. L., & Adams, C. M. (Eds.). (1986). *Linguistic Atlas of the Gulf States* (Vols. 1–7). Athens, Georgia: University of Georgia Press.
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., Swingley, 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.
Rasico, P. D. (1993). Dues questions de fonologia diacrònica catalana. In P. D. Rasico, *Estudis i documents de lingüística històrica catalana* (pp. 9–19). Barcelona: Curial.
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. (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. (2016). Phonetic Distinctiveness vs. Lexical Contrastiveness in Non-Robust Phonemic Contrasts. *Laboratory Phonology: Journal of the Association for Laboratory Phonology*, *7*(1), 1–29. https://doi.org/10.5334/labphon.17
Renwick, M. E. L., Vasilescu, I., Dutrey, C., Lamel, L., & Vieru, B. (2016). Marginal Contrast Among Romanian Vowels: Evidence from ASR and Functional Load. In *Proceedings of Interspeech 2016* (pp. 2433–2437). San Francisco, CA. Retrieved from http://www.isca-speech.org/archive/Interspeech_2016/pdfs/0762.PDF
Schilling, N. (2014). Surveys and interviews. In R. J. Podesva & D. Sharma (Eds.), *Research Methods in Linguistics* (pp. 96–115). Cambridge: Cambridge University Press.
Schütze, C. T. (1996). *The empirical base of linguistics*. Chicago: University of Chicago Press.
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.

Shitara, Y. (1993). A survey of American pronunciation preferences. *Speech Hearing and Language*, *7*, 201–232.
Sim, J., & Wright, C. C. (2005). The kappa statistic in reliability studies: use, interpretation, and sample size requirements. *Physical Therapy*, *85*(3), 257–268.
Simonet, M. (2014). Phonetic consequences of dynamic cross-linguistic interference in proficient bilinguals. *Journal of Phonetics*, *43*, 26–37.
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.
Tilsen, S., & Cohn, A. C. (2016). Shared Representations Underlie Metaphonological Judgments and Speech Motor Control. *Laboratory Phonology: Journal of the Association for Laboratory Phonology*, *7(1)*(14), 1–33. https://doi.org/http://doi.org/10.5334/labphon.52
Trubetzkoy, N. S. (1939). *Principles of Phonology*. (C. Baltaxe, Trans.). Berkeley: University of California Press.
Walker, A., García, C., Cortés, Y., & Campbell-Kibler, K. (2014). Comparing social meanings across listener and speaker groups: The indexical field of Spanish /s/. *Language Variation and Change*, *26*(02), 169–189. https://doi.org/10.1017/S0954394514000088
Wasow, T., & Arnold, J. (2005). Intuitions in linguistic argumentation. *Lingua*, *115*(11), 1481–1496. https://doi.org/10.1016/j.lingua.2004.07.001
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.
Wells, J. C. (1990). *Longman Pronunciation Dictionary* (First Edition). Harlow: Longman.
Wells, J. C. (1999a). British English pronunciation preferences: a changing scene. *Journal of the International Phonetic Association*, *29*(01), 33–50.
Wells, J. C. (1999b). Pronunciation preferences in British English: a new survey. In *Proceedings of the 14th International Congress of Phonetic Sciences* (pp. 1–4). San Francisco.
Wells, J. C. (2000). *Longman Pronunciation Dictionary* (Second Edition). Harlow: Pearson Education Limited.
Wells, J. C. (2003). Pronunciation research by written questionnaire. In *Proceedings of the 15th International Congress of Phonetic Sciences (ICPhS 2003)* (pp. 215–218). Barcelona.
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.