
Detailing the impact of social variables on the production of the Catalan mid-vowel contrasts by early Spanish-Catalan bilinguals

Zoi Kotsoni ¹  0000-0003-1724-3749

¹ Universitat Autònoma de Barcelona (Spain)

DOI: 10.1344/efe-2024-33-57-68

Corresponding address: zoe.kotsoni@outlook.com

Received: 06/06/2023 Accepted: 15/01/2024 Published: 06/03/2024

Kotsoni, Z. (2024). Detailing the impact of social variables on the production of the Catalan mid-vowel contrasts by early Spanish-Catalan bilinguals. *Estudios de Fonética Experimental*, 33, 57–68. <https://doi.org/10.1344/efe-2024-33-57-68>



Subject to the license Creative Commons CC BY-NC-ND 4.0 DEED
Attribution-NonCommercial-NoDerivs 4.0 International

ABSTRACT

The present study investigates variability in the production of Catalan vowels by Barcelona young, middle-aged, and older adults who speak the Central Catalan variety. The degree of exposure to and use of Central Catalan varies among speakers as half of the participants are second-generation and subsequent-generation migrants from other regions of Spain, where the vernacular is Spanish. All speakers have been born, raised, and schooled in Barcelona, and have acquired both Central Catalan and Spanish. Central Catalan possesses two sets of phonemic mid-vowels (/e/–/ɛ/ and /o/–/ɔ/), unlike Spanish which has a single vowel per set (/e/ and /o/). This study aims to detail the Catalan mid-front and mid-back vowel contrasts used by speakers of different gender, age, language use, and exposure to Catalan from the bilingual speech community of Barcelona.

KEYWORDS

speech production; language contact; variability in production; mid-vowel contrasts; language-dominant group

Analitzant l'impacte de les variables socials en la producció dels contrastos de vocals mitjanes catalanes dels primers bilingües castellà-català

RESUM

Aquest estudi investiga la variabilitat en la producció de vocals catalanes per part d'adults barcelonins, joves, de mitjana edat i grans, parlants de català central. El grau d'exposició i ús del català central varia entre els parlants, ja que la meitat dels participants són migrants de segona generació i de generacions posteriors d'altres regions d'Espanya, en què l'idioma vernacle és el castellà. Tots els parlants han nascut, crescut i escolaritzat a Barcelona, i han adquirit tant el català central com el castellà. El català central posseeix dos conjunts de vocals mitjanes fonèmiques (/e-/ɛ/ i /o-/ɔ/), a diferència del castellà, que té una sola vocal per conjunt (/e/ i /o/). Aquest estudi té com a objectiu detallar els contrastos vocàlics mitjans-anteriors i mitjans-posteriors catalans utilitzats per parlants de diferent gènere, edat, ús lingüístic i exposició al català de la comunitat de parla bilingüe de Barcelona.

MOTS CLAU

producció de la parla; contacte lingüístic; variabilitat en la producció; contrastos de vocals mitjanes; grup de llengua dominant

1. Introduction

The Central Catalan variety spoken in the province of Barcelona, the eastern half of Tarragona province, and into the entire province of Girona has been classified in the eastern dialectal block (Milà i Fontanals, 1861). Central Catalan has a four-height vowel system in which the close-mid front /e/ in tandem with the open-mid front /ɛ/ constitute a contrastive vowel pair in stressed syllables. Similarly, a minimal vowel contrast is retained between the close-mid back /o/ and the open-mid back /ɔ/ vowels (see Table 1). Due to the Central Catalan vowel reduction rule referenced in *Gramàtica del català contemporani* (Mascaró, 2002, vol. 1, pp. 92–93), the four-height stressed vowel system results in a three-height unstressed vowel system with a schwa and two high vowels. In particular, the three-way phonemic contrast between /e ε a/ is neutralised in [ə] in unstressed syllables, while /o/ and /ɔ/ reduce to [u] (see Table 2).

height 4	i	u
height 3	e	o
height 2	ɛ	ɔ
height 1	a	

Table 1. The four-height stressed vowel system of Central Catalan.

height 2	i	u
height 1	ə	

Table 2. The two-height unstressed vowel system of Central Catalan.

The Spanish vowel system has a single degree of aperture for the mid-front and mid-back vowels in which /e/ and /o/ share the second-height position. The third-height position is occupied by the /i/ and /u/ high vowels and the first height is solely occupied by /a/. Table 3 below illustrates the three-height vowel system of Spanish having a five-member vowel inventory, whose number is inalterable in unstressed syllables. While vowel reduction does not seem to occur in Spanish, Martínez Celdrán & Fernández Planas (2007, p. 188) based on an

articulatory analysis of their data, demonstrated that the openness of the Spanish mid-vowels /e/ and /o/ occurs in a context-related phonetic environment. Consequently, /e/ and /o/ have open variants when they are in contact with [r], before [x], in a syllable closed by any consonant, with the constraint of /e/, which remains closed in syllables surrounded by [m, n, s, d, θ] and in rising diphthongs with a palatal glide.

height 3	i	u
height 2	e	o
height 1	a	

Table 3. The three-height vowel system of Spanish.

The discussion in this paper of Catalan stressed mid-vowels is confined to a specific linguistic factor, known as *inherent variation*, which is “an internally generated variability that happens when the vowel inventory size and quality characteristics alone are responsible for the maintenance or loss of an audible difference between vowels” (Kotsoni, 2019, p. 10). The analysis can further be extended to cover broader categories of externally motivated variation which, along with inherent variation, serve to develop an overall comparative picture. *External variation* is not conceived as directly relevant to the Catalan vowel inventory; it is determined by social variables, such as gender, age, and language-dominant group, and by social constructs developed within society, such as speakers’ attitudes and exposure to the languages under study, the intensity of contact, language use in childhood and adulthood and the degree of speakers’ integration within a linguistic system.

Some variables mentioned above have relevance for externally motivated variation in speech. This study discusses the effect of gender, age, and language-dominant group on the mid-vowel pairs by considering the background of each bilingual participant on the grounds of their social characteristics.

Gender has played a significant role in variationist sociolinguistic studies, as a link can be readily noticed between females/males and standard/non-

standard language use. Scholars have been investigating gender as being a trigger for variation in language use through evidence-based studies. The academic research and debate, vivid since 1922,¹ has reached the following conclusions: “In careful speech women use fewer stigmatized forms than men, and are more sensitive than men to the prestige pattern” (Labov, 1972, p. 243). “Females show a greater sensibility to socially evaluative linguistic forms than do males” (Wolfram, 1969, p. 78), “Women on average deviate less from the prestige standard than men” (Coates & Cameron, 1989, p. 13).² These conclusions highlight that women’s use of language is perceived to carry more prestige or to be closer to the standard variety than men’s use of language.

Age is distinct from other social factors in certain key ways. Age grading, for example, has been correlated with variation in speech in a way that makes it more possible to reflect an age-graded change than a change in progress. Tagliamonte (2012, p. 44) argues that when linguistic features become evident in apparent-time and show age differences, they can point toward the phenomenon of *age grading*. Certain types of speech may at any time be commoner in young-aged speakers, e.g., slang in schoolchildren and students than in adults. Age-graded change can best be understood regarding people of different ages who “use language differently simply because they are at different stages in their life” (Tagliamonte, 2012, p. 47), and they use “speech appropriate to their age group” (Wardough, 2002, p. 194). I also focus here on the definition of *age grading* provided by Labov: “If individuals change their linguistic behaviour throughout their lifetimes, but the community as a whole does not

change, the pattern can be characterized as one of age grading” (Labov, 1994, p. 84). Differences in the population sample chosen for participation in this study are patterned among the speech of individuals of 15–34 years, 35–54 years, and 55–75 years, bringing generation differences into the analysis. It is possible, thus, to study how certain phonemes, such as the open-mid vowels occur in an age group but miss from another at a particular place and time.

Scholars (Montrul, 2016; Van Coetsem, 1988) have studied the *language dominance* used in either linguistic, social, or both contexts, so this notion comes within an interdisciplinary scope. Here I review some previous work comprising the study of language dominance in a linguistic context to illustrate several dimensions: a) linguistic proficiency; b) input; and c) context and use (Montrul, 2016, p. 16). *Language dominance* first consists of the internal component of linguistic proficiency, the fact of having the skill and experience to master language(s). Second, input as an external component reflects the sheer amount of language(s) a speaker is exposed to in aspects of grammar and vocabulary development. A third functional component of the language dominance factor is the context in which language(s) are spoken along with the way speakers use them to serve their communication needs. In the bilingual language community of Barcelona, second and later generations of internal migrants, who have acquired Central Catalan through the Catalan immersion system may become dominant in Catalan, unlike internal migrants who have not received their primary and secondary education in Catalan but acquired Catalan as adult learners. However, even second or later generations of internal migrants may

¹ Jespersen (1922, p. 247) argued that women, in comparison to men, demonstrated conservatism in their linguistic behaviour, whereas men seemed to be far more innovative in language using ‘new’ and ‘fresh expressions’.

² Note that sociolinguists have repeatedly established theories on gender (or sex) which has always been subject to academic debates. Here, I summarise the following views concerned with the study of gender as this emerges through participants’ observation in previous sociolinguistic fieldwork. ‘Females show more awareness of prestige norms in both their actual speech and attitudes towards speech’ (Wolfram & Fasold, 1974, p. 93). “Women, allowing for

other variables such as age, education, and social class, produce on average linguistic forms which more closely approach those of the standard language or have higher prestige than those produced by men” (Trudgill, 1983, p. 161). Thus, gender, as a key social variable has been treated in a way that shifted the focus of women’s conservatism in language use towards the idea that stylistic variation, namely, alternation of discourse structure to match the context and situation, lags behind women’s use of fewer stigmatized and non-standard forms in comparison to men’s language use.

also retain some regional variants to show their allegiance to particular geographical groups (Evans & Iverson, 2007; Foulkes & Docherty, 1999; Trudgill, 1986).

This study begins with the observation of the Central Catalan vernacular, as speakers even unconsciously may opt for using the Catalan close-mid vowels /e/ and /o/ in utterances where an /ɛ/ or /ɔ/ are required. Straightforwardly, the present paper focuses on variation in the production of the Catalan open-mid vowels, variants that have been found to appear less frequently than the close-mid vowels in languages spoken around the world. Broadly, it deals with the phonetic effects of linguistic interaction in the city of Barcelona where language contact occurs between Catalan and Spanish which are genetically related. Of all the variables this study surveys, the presence, or absence of an audible difference between the mid-vowel pairs, the effects of social variables, gender, age, and language dominance are restricted to language contact explanations. Language contact is a widespread phenomenon across bilingual or multilingual communities. The literature is departing from field research done in natural settings where the speech community is entirely monolingual (Labov, 1966; Milroy, 1980; Trudgill, 1974) and moving toward works on the dialectal configuration of a majority language spoken by natives in a bilingual community (see Blas Arroyo, 2011, for a review of Spanish in contact with Catalan).

Contact can influence any of the two languages (Catalan, Spanish) involved. In the context of this study, however, the important question concerns the extent to which the contact that speakers of the Central Catalan variety have had, with speakers of Spanish can influence the vowel quality of /ɛ/ and /ɔ/. The present paper as well as similar fieldwork conducted in Catalan-speaking areas that may lead to different outcomes of language contact between Catalan and Spanish, ranging from limited *mergers* to the maintenance of the mid-vowel contrasts, aims to account for variation in the production of the mid-vowels in Catalan, and concern the degree of influence of Spanish, which is embedded in society.

In this section, I will summarize studies within experimental phonology that have had two main characteristics in common. First, researchers have used a sociolinguistic methodology to collect data from Catalan-Spanish bilingual communities and included acoustic analysis to support their proposal of sound variation in the mid-vowel contrasts of Catalan. Second, all these studies have a great deal in common with the present paper in determining whether a merger or near merger is the result of contact, although it would not be possible for these similar bilingual communities (Barcelona, Majorca) to be affected equally by contact between Catalan and Spanish. It would be fair to mention that most studies have focused on one mid-vowel contrast; either to the mid-front or to the mid-back Catalan vowel contrast. Cortés et al. (2019) conducted a careful study on *ongoing changes* in the mid-front vowels of the Catalan language spoken in three districts (*Gràcia*, *Eixample*, and *Nou Barris*) of the city of Barcelona. The maintenance of the mid-front vowel contrast was found in the sample with the highest degree of Catalan presence (*Gràcia* and *Eixample*), while the children who live in *Nou Barris* (a district with a high degree of Spanish presence) merged between /e/ and /ɛ/. It is interesting to note that the loss of the mid-front vowel contrast in Catalan correlates with age and social setting and hence can be used as an example of fundamental difference across generations. A recent experimental study analysed the realisation of the stressed Catalan /o/–/ɔ/ and the Spanish /o/ in the Catalan-Spanish contact varieties of the island of Majorca (Simonet, 2011), to determine how interaction patterns classified by dominant and non-dominant phonetic subsystems have been developed into each language-dominant group. As concerns his analysis results, Catalan-dominant speakers produced a Catalan /o/–/ɔ/ that had no equivalent in Spanish. Spanish-dominant speakers, however, developed a separate vowel category, such as one for their Spanish /o/ and one for their merged Catalan /o/ + /ɔ/ to show that their mid-back vowel system in Catalan is distinct from their mid-back vowel system in Spanish. Similarly, Amengual (2014) explored qualitative differences in the context of two sets of data as these are found in the speech of Spanish-speaking and Catalan-

speaking dominant bilinguals who reside on the island of Majorca. The author found a correlation between performance on production and perception tasks for mid-vowel contrasts in Majorcan Catalan. Catalan-Spanish bilinguals who produced the mid-vowels with smaller Euclidean distances tended to present a higher error rate in the perception task than bilinguals who produced more robust mid-vowel contrasts.

2. Method

The present paper treats the Barcelona metropolitan area as a relatively homogenous dialect area where, regarding Catalan, a single language variety is manifested by residents in several ways; used daily within the education system, through media, and in the administrative services which can also be drafted in Spanish. The Barcelona Catalan dataset is composed of bilingual subjects in Catalan and Spanish whose language performance is representative of the vernacular of the city of Barcelona in Spain.

2.1. Participants

A randomly stratified sampling procedure with social networking was used in this study to collect data from Spanish/Catalan bilingual participants. Seventy-two subjects from two language backgrounds were recorded: thirty-six Catalan-dominant bilingual speakers and thirty-six Spanish-dominant bilingual speakers, aged 15–75 years, and divided among young, middle-aged, and older adults. The individuals in each age subgroup did not outrate or were less than twelve. Besides a perfect balance has been kept between Catalan-dominant and Spanish-dominant bilinguals across their three different age groups ($n = 12$), and the selection of individuals based on their gender has been done randomly.

All subjects were recruited from the Barcelona metropolitan area and have been born and raised in Barcelona, even though not all of them have been fully educated in Catalan. Participants of the Catalan-dominant bilingual group had a mean age of 43.7 years ($SD = 16.9$ years) and participants of the Spanish-dominant bilingual group had a mean age

of 42.6 ($SD = 18.3$ years) at the time of recruitment. All speakers were exposed to prolonged bilingualism on a daily basis, but their language use differed according to the conditions that they lived or worked in and the way that these influenced how they felt or how effectively they could communicate in Catalan and Spanish.

2.2. Stimuli and apparatus

The stimuli consisted of 12 words in the carrier sentences:

Faltaven uns minuts per a un quart de **set** però ell es comportava com si **faltés** una eternitat [...] a un quart de **set** de la tarda [...] treballava al restaurant “**Set Portes**”, un dels més emblemàtics de la ciutat amb més de **dos**-cents anys d’antiguitat [...] a pocs metres per **sota** de la plaça de la Reforma [...] al capdamunt de la pila de diaris de **sota** la taula. El talent és un **dot** molt estimat [...] el seu pare desitjava més que **completés** els seus estudis amb una carrera universitària [...] volia convertir-se en un **sacerdot** [...] que fóra millor que ella aportés en **dot** sensibilitat, més que no pas que **aportés** en dot riquesa material.

The test words were elicited using an authentic Catalan passage. All target sounds were elicited in word-final (WF) position, e.g., *dot*, *set*, *faltés*, *completés*, *aportés*, and *sacerdot*, except for /o/ where the target sound was found in both word-initial (WI) and word-final (WF) position, e.g., /*sota*/ and /*dos*/. All Catalan stressed vowels were produced within the same CVC frame, preceded by /d/, /s/, and /t/ and followed by /s/ and /t/. The C_C context consisted of the voiced /d/ and voiceless /t/, /s/ alveolars to minimise possible effects of a different consonant environment on vowel formants. All recordings were made with a portable digital recorder (WS-550M) with a built-in stereo microphone.

2.3. Procedure

Throughout this section, I describe how I proceeded with a three-stage interview to achieve a random sample of 72 bilingual individuals, divided among 36 Catalan-dominant and 36 Spanish-dominant bilinguals. Interviews were conducted in Catalan.

Sp.	G.	Age grad.	Catalan general score	Spanish general score	Domin. score
NP	F	15-34	167.54	129.85	37.69
CL	F	15-34	172.17	126.12	46.05
GA	F	15-34	169.81	129.57	40.24
SE	F	15-34	169.27	148.94	20.33
IA	F	15-34	177.26	127.31	49.90
EM	F	15-34	158.55	155.37	3.18
JP	F	15-34	122.23	116.87	5.39
HP	M	15-34	168.09	136.48	31.61
AV	M	15-34	171.72	112.05	59.67
DR	M	15-34	166.27	132.67	33.60
JVA	M	15-34	156.86	94.81	62.05
ACC	M	15-34	175.62	144.92	30.70
PC	F	35-54	185.34	131.76	53.58
S	F	35-54	173.45	118.50	54.95
CE	F	35-54	154.38	102.62	51.76
RM	F	35-54	165.54	150.46	15.08
AT	F	35-54	177.17	132.30	44.87
GCM	F	35-54	148.93	115.78	33.15
AA	F	35-54	158.10	158.10	16.25
JS	F	35-54	185.07	149.83	35.24
JV	M	35-54	140.85	123.86	16.99
A	M	35-54	139.84	117.05	22.79
AC	M	35-54	203.56	111.52	92.04
MO	M	35-54	146.93	145.47	1.46
YS	F	55-75	164.18	127.12	61.58
MT	F	55-75	167.63	112.96	37.06
RB	F	55-75	154.28	114.60	54.67
CCB	F	55-75	166.09	66.54	39.68
MF	F	55-75	162.46	120.68	99.56
PM	F	55-75	158.01	96.25	41.78
MP	F	55-75	143.75	108.69	61.76
RMC	F	55-75	164.73	96.34	35.06
FM	M	55-75	132.67	81.86	68.39
JN	M	55-75	152.65	115.23	50.81
JPC	M	55-75	161.64	119.34	37.42
JF	M	55-75	174.72	119.04	42.30

Table 1. Overall distribution of global language scores and dominance score, Catalan-dominant bilinguals.
Sp. = Speaker; G. = Gender.

Sp.	G.	Age grad.	Spanish general score	Catalan general score	Domin. score
AH	F	15-34	174.81	67.19	107.62
ER	F	15-34	160.55	143.2	17.35
MC	F	15-34	164.36	139.08	25.28
JA	F	15-34	159.82	122.00	37.42
BG	F	15-34	136.21	102.97	33.24
IR	F	15-34	107.39	39.62	67.77
PCR	F	15-34	151.31	133.94	17.37
SA	F	15-34	177.62	108.24	69.39
JB	M	15-34	174.72	127.03	47.68
JPE	M	15-34	168.54	134.48	34.06
HC	M	15-34	114.23	175.26	61.03
DM	M	15-34	180.89	107.87	73.02
OT	F	35-54	193.24	122.04	71.21
MV	F	35-54	179.35	128.31	51.04
GC	F	35-54	191.24	157.45	33.79
MA	F	35-54	174.99	127.40	44.59
SR	F	35-54	138.31	117.04	22.27
SL	F	35-54	143.34	97.01	46.33
EL	F	35-54	114.98	59.02	55.97
SM	F	35-54	144.11	136.94	7.17
MM	F	35-54	117.41	109.96	7.45
LH	M	35-54	184.34	54.93	129.41
JE	M	35-54	173.18	117.86	55.31
OM	M	35-54	158.65	97.07	61.58
AG	F	55-75	186.52	94.07	92.45
AZ	F	55-75	248.00	44.00	204.00
AM	F	55-75	171.90	90.38	81.52
EG	F	55-75	154.65	132.30	22.35
EB	F	55-75	159.10	118.77	40.32
MCC	F	55-75	186.79	73.37	113.43
TB	F	55-75	162.00	66.75	95.25
PT	F	55-75	164.36	126.22	38.14
EP	M	55-75	148.29	74.18	74.11
AMP	M	55-75	136.31	62.11	74.20
JAM	M	55-75	185.71	45.40	140.31
J	M	55-75	140.75	133.49	7.26

Table 2. Overall distribution of global language scores and dominance score, Spanish-dominant bilinguals.
Sp. = Speaker; G. = Gender.

Given the fact that a stranger arouses the level of attention paid to speech, participants first answered an array of Frequently Asked Questions (FAQs) related to the following topic categories; *a*) demography and languages spoken with family and friends; *b*) education; *c*) hobbies and family; and *d*) bilingualism. Focusing on a particular aspect or angle of each broad category, helped participants pay less attention to speech. Second, subjects were instructed to read aloud an authentic Catalan passage with the words that included the target sounds. The Catalan passage has been read aloud twice to circumvent orthographic effects that might have arisen in the first instance. Third, all participants completed the Bilingual Language Profile (BLP) detailed questionnaire, adapted from Birdsong et al. (2012), which consisted of questions on four different modules: language history, language use, language proficiency, and language attitudes. Based on speakers' feedback on the relevant modules, each participant has ascribed a dominant language. The BLP was available for participants to fill out in two language versions (Spanish or Catalan) and was provided in their language of preference.

The validity of the BLP instrument in analyses based on 72 participants has been demonstrated following a manual scoring procedure. Question responses in the BLP modules were each associated with a number on a scale, which ranged from 0 to 20. I then multiplied the score for each module by a factor of *a*) language history $\times 0.454$, *b*) language use $\times 1.09$, *c*) language proficiency $\times 2.27$, and *d*) language attitudes $\times 2.27$, as suggested by Birdsong et al. (2012). Adding all module totals together (54.48 + 54.5 + 54.48 + 54.48) to yield a global score for each language, the highest total point value possible is 217.94. One language total was subtracted by the other to provide a dominance score. No scores of zero that would indicate perfectly balanced bilingualism have been noticed.

Table 1 provides a direct comparison among the 15–34, 35–54, and 55–75-year-old speakers which reveals balanced bilingualism in young and middle-aged Catalan-dominant bilinguals, whose dominance score was as close as possible to zero (see

speakers EM, JP, and MO below for their dominance score near zero).

Table 2 reveals several cases of balanced bilingualism in middle-aged and older Spanish-dominant bilinguals (see speakers SM, MM, and J below for their dominance score near zero).

2.3.1. Acoustic analysis

F1 and F2 were measured for the mid-vowels in each target word read in the carrier sentence, giving two sets of measurements for each word. The target words consisted of 12. Each target word was recorded twice, as a result, each participant produced (12 \times 2) 24 experimental tokens. F1 and F2 were averaged across these repetitions per word and all measurements were used in the subsequent analyses. The onsets and offsets of each vowel were located manually by looking at the waveform, spectrogram, and intensity curve. F1 and F2 were extracted using hand-corrected LPC analyses. Formant frequencies were measured from the midpoint of the steady-state portion of the vowel. F1 and F2 were measured using *Praat* (Boersma & Weenink, 1992–2022). The intensity settings in the spectrogram were configured at 30 (dB), which helped to eliminate soft background noises.

The data was submitted to a speaker normalisation procedure to minimise or eliminate acoustic variation in vowel measurements due to physiological differences (Fabricius, 2008). The following formula introduced by Lobanov (1971) was used to normalise the acoustic measurements: $F_n[V]^N = (F_n[V] - MEAN_n) / S_n$.

2.3.2. Statistical analysis

A linear mixed effects model was built, which included GROUP (Catalan dom., Spanish dom.), GENDER (female, male), AGE (young, middle-aged, older), and VOWEL (front, back) as fixed factors, and participant as a random factor. Post-hoc tests were run wherever a given factor showed up as significant in the model, to elucidate specific differences among the levels of that variable.

3. Results

The results demonstrated that Spanish bilingual speakers have closer mid-vowel productions of the Catalan /e/–/ɛ/ and /o/–/ɔ/ phonological contrasts, whereas Catalan bilingual speakers retain with a larger acoustic distance the split between the /e/–/ɛ/ and /o/–/ɔ/. Young Spanish bilingual speakers tend to modify their production of the mid-open Catalan vowel [ɔ] to a vowel phoneme that occupies an intermediate position between the Catalan /o/–/ɔ/ phonetic contrast. All speakers were found to distinguish better between the mid-front vowels than between the mid-back vowels. The results further demonstrated that the Catalan female bilingual speakers have the most robust mid-vowel

differentiation among participants. Implications that suggest that variation is involved in this study are discussed below.

3.1. Euclidean distances between mid-vowels within each pair

Figures 1 and 2 show boxplots of the data to provide an exploratory data analysis based on visual displays which should highlight patterns that help us summarize and compare groups of data effectively. The boxplots in both figures provide visual displays describing the distribution of the front and back mid-vowel contrasts by group (Catalan-dominant, Spanish-dominant) and age (young, middle-aged, and older adults).

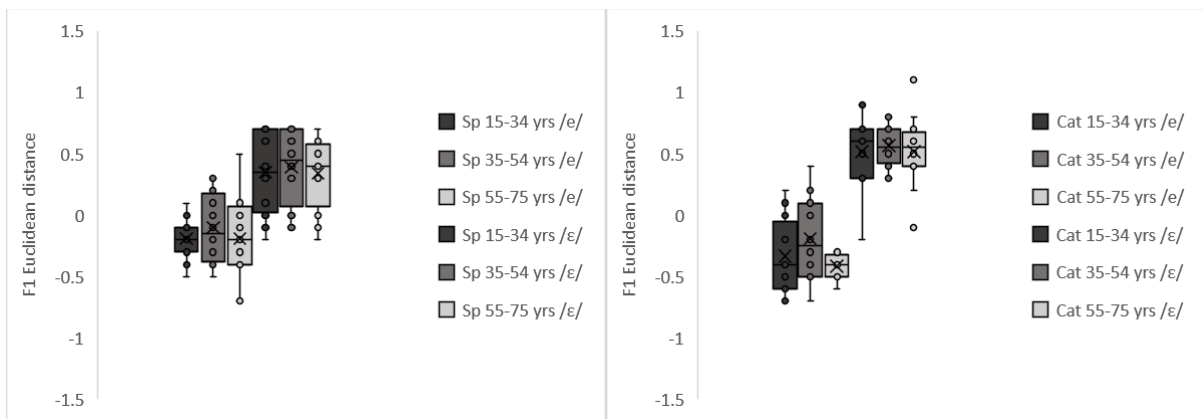


Figure 1. Boxplots of the Euclidean distance for mid-front vowels across age groups of Spanish and Catalan-dominant bilinguals.

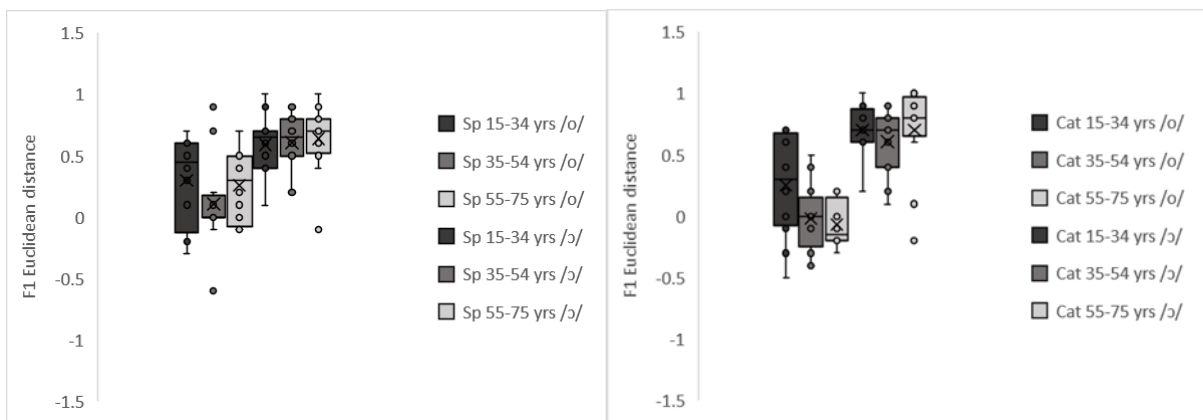


Figure 2. Boxplots of the Euclidean distance for mid-back vowels across age groups of Spanish and Catalan-dominant bilinguals.

As displayed in Figure 1, Spanish-dominant bilinguals (SPs) 15–34 years seem to have a remarkably narrow dispersion for /e/, and likewise Catalan-dominant bilinguals (CATs) 55–75 years seem to present an implausibly narrow dispersion for /e/. SPs from 15 to 75 years appear to present a greater dispersion for /ε/ than CATs from 15 to 75 years. The three Catalan age groups have a greater split than the three Spanish age groups between the mid-front vowels.

As displayed in Figure 2, both SPs and CATs 15–34 years appear to present a greater dispersion for /o/ but a narrower for /ɔ/, even not as narrow as that of SPs 35–54 years for /o/. CATs have the greatest split between the mid-back vowels, whereas SPs 15–34 years appear to barely merge the mid-back vowels and should have one category.

3.2. Formant frequencies of mid vowels

F1 and F2 values are displayed in Figure 3. The young, middle-aged, and older Catalan-dominant speakers provide more robust vowel contrasts in their /e/–/ε/ and /o/–/ɔ/ productions than the Spanish-dominant speakers. All speakers produce a larger acoustic distinction between the mid-front vowels. While the mid-back vowels are closer for both CATs and SPs, SPs 15–34 years appear to barely merge their two-category mid-back vowel /o/–/ɔ/ and to be close to possessing one.

3.2.1. Mid-front vowels

There was a main effect of GROUP for F1 [$F(1, 67) = 12.59, p < .001$], and a significant interaction VOWEL × GROUP for F1 [$F(1, 62) = 10.97, p = .0015$]. Inspection of the data revealed that the Catalan-dominant group produced /e/–/ε/ with greater F1 mean distance than did the Spanish-dominant group resulting in a two-category mid-front vowel system. There was no main effect of GENDER, AGE, or VOWEL for F1 ($p > 0.05$), as well as no other statistically significant interaction.

3.2.2. Mid-back vowels

There was no main effect of GROUP, GENDER, AGE, or VOWEL ($p > 0.05$), confirming that speakers produced /o/ and /ɔ/ with a similar F1. There was a significant interaction GENDER × GROUP for F1 [$F(1, 66) = 7.46, p = .008$]. Pairwise comparison tests conducted on the data revealed that this was associated with a larger difference for F1 [$F(1, 66) = 4.17, p = .045$], confirming that the Catalan-dominant females had a greater mid-back vowel split than that of Spanish-dominant females. In parallel to these production results, no significant interaction emerged between Catalan-dominant and Spanish-dominant males for F1 ($p > 0.05$). Pairwise comparisons also yielded a significant difference between Catalan-dominant speakers of different genders [$F(1, 66) = 9.94, p = .002$], indicating that females

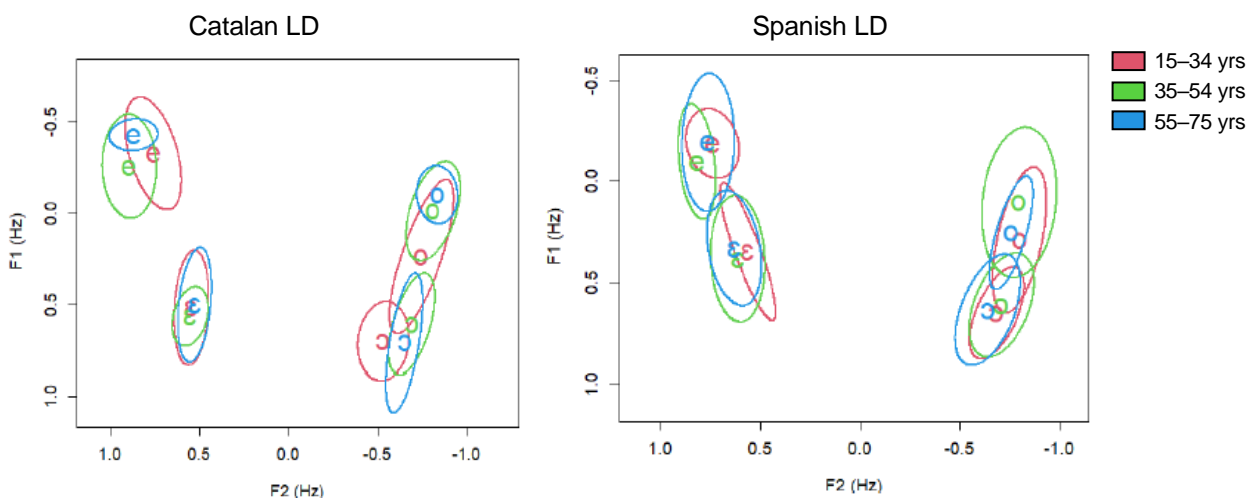


Figure 3. F1 and F2 plots for speakers' productions of target Catalan mid vowels split by age.

performed better than males. The F1 difference between Spanish-dominant females and males did not achieve any significance.

The results demonstrated that language dominance turned up as significant; CATs had a greater mid-front vowel contrast than SPs. Even though gender did not show up as significant in the omnibus model, there was evidence for systematic variability in production that is associated with gender and language dominance as the mid-back vowel contrast turned out to be far better attained by Catalan-dominant females.

Although participants were selected on the basis of being from a homogeneous background (they have all been born, raised, and attended school in Barcelona), differences in production emerge as an overall result of language dominance differentiation, rather than being reflective of gender-based differences that are found between the levels (female, male) of this category. What the results suggest to a certain extent is that Catalan females perform better than Catalan males, which is in line with previous conclusions reached by scholars (see Labov, 1972; Wolfram, 1969; Coates & Cameron, 1989). A generalisation that can be made from the post hoc tests performed in this study is that Catalan women's use of language is perceived to carry more prestige or to be closer to the standard variety than Catalan men's use of language.

As expected, Catalan bilingual speakers tended to retain the split between /e/–/ɛ/ and /o/–/ɔ/ as a result of attending school in Catalan, particularly for young and middle-aged Catalan speakers who have attended courses entirely taught in Catalan and have used their first language (L1-Catalan) in the classroom. Older Catalan speakers tended to produce similar mid-vowel contrasts when speaking Central Catalan as a result of using L1 in the family.

4. Discussion

Language dominance and age have been some common social factors discussed above, which have relevance for variability in production and set the stage for similar linguistic outcomes. In line with a previous study of Catalan and Spanish groups and their productions of the mid-back vowel phonemes (Simonet, 2011), the present results suggest that the young Spanish bilingual speakers use a closer mid-back vowel category to produce [ɔ] when speaking Catalan. The value of linguistic performance has been illustrated with a pattern involving a closer mid-back phonetic category in Catalan associated with Spanish-dominant bilingual speakers. In analysing more linguistic effects of variability in speech production, age should be considered as a highly relevant social factor in explaining externally motivated variation. This point is compatible with the results of both Cortés et al. (2019) paper and our study: Spanish bilingual children showed an advanced merger between /ɛ/ and /e/, whereas the young Spanish bilingual speakers (15–34 years) displayed an in-between /o/ and /ɔ/ vowel phoneme.

Our results further demonstrated that the Catalan female bilingual speakers have the most robust mid-vowel differentiation among participants. Our observations suggest that women in this bilingual speech community when making use of their L1-Catalan are more inclined than men to indulge in careful speech and, consequently, in the maintenance of the mid-vowel contrasts. A good deal of research remains to be done since language dominance combined with gender (Catalan-dominant female bilinguals) shows evidence for the maintenance of the mid-vowel contrasts. Further research is required, in particular since a pattern involving a closer mid-vowel phonetic category in Catalan has been associated with the value of social setting, which appears to be less effective than the value of schooling in L2-Catalan in the Spanish-speaking districts of Barcelona.

References

- Amengual, M. (2014). The perception and production of language-specific mid-vowel contrasts: Shifting the focus to the bilingual individual in early language input conditions. *International Journal of Bilingualism*, 20(2), 1–21. <https://doi.org/10.1177/1367006914544988>
- Birdsong, D., Gertken, L. M., & Amengual, M. (2012, January 20). *Bilingual Language Profile: An easy-to-use instrument to assess bilingualism*. COERLL. <https://sites.la.utexas.edu/bilingual/>
- Blas Arroyo, J. L. (2011). Spanish in contact with Catalan. In M. Díaz-Campos (Ed.), *The Handbook of Hispanic Sociolinguistics* (pp. 374–394). Wiley-Blackwell. <https://doi.org/10.1002/9781444393446.ch18>
- Boersma, P., & Weenink, D. (1992–2014). Praat: Doing phonetics by computer (Version 5.3.82) [Computer program]. <http://www.praat.org/>
- Coates, J., & Cameron, D. (Eds.). (1989). *Women in their speech communities*. Routledge.
- Cortés, S., Lleó, C., & Benet, A. (2019). Weighing factors responsible for the production of the Catalan vowel /ɛ/ versus /e/ contrast in three districts of Barcelona. *International Journal of Bilingualism*, 23(6), 1264–1277. <https://doi.org/10.1177/1367006918781058>
- Evans, B. G., & Iverson, P. (2007). Plasticity in vowel perception and production: A study of accent change in young adults. *The Journal of the Acoustical Society of America*, 121(6), 3814–3826. <https://doi.org/10.1121/1.2722209>
- Fabricius, A. H. (2008, September 16). *Vowel normalisation in sociophonetics: When, why, how?* [Conference presentation]. Sociolinguistics Circle, Copenhagen University, Copenhagen.
- Foulkes, P., & Docherty, G. J. (1999). Urban Voices – Overview. In P. Foulkes, & G. J. Docherty (Eds.), *Urban Voices: Accent Studies in the British Isles* (pp. 1–24). Arnold.
- Jespersen, O. (1922). *Language: Its nature, development and origin*. Allen and Unwin.
- Kotsoni, Z. (2019). *A sociophonetic analysis of the production of mid-vowel contrasts in Catalan spoken in Barcelona* [Doctoral dissertation, Universitat Autònoma de Barcelona]. Tesis Doctorals en Xarxa. <https://hdl.handle.net/10803/669724>
- Labov, W. (1966). *The social stratification of English in New York City*. Center for Applied Linguistics. <https://doi.org/10.1017/CBO9780511618208>
- Labov, W. (1972). Some principles of linguistic methodology. *Language in Society*, 1(1), 97–120. <https://doi.org/10.1017/S0047404500006576>
- Labov, W. (1994). *Principles of linguistic change, Vol. 1: Internal factors*. Blackwell.
- Lobanov, B. M. (1971). Classification of Russian vowels spoken by different speakers. *Journal of the American Statistical Association*, 49(2), 606–608. <https://doi.org/10.1121/1.1912396>
- Martínez Celdrán, E., & Fernández Planas, A. M. (2007). *Manual de fonética española: Articulaciones y sonidos del español*. Ariel.
- Mascaró, J. (2002). El sistema vocàlic. Reducció vocàlica. In J. Solà, M.-R. Lloret, J. Mascaró, & M. Pérez Saldanya (Eds.), *Gramàtica del català contemporani* (pp. 89–123). Empúries.
- Milà i Fontanals, M. (1861). *De los trovadores en España. Estudio de la lengua y poesía provenzal*. Librería de Joaquín Verdager.
- Milroy, L. (1980). *Language and social networks*. University Park Press.
- Montrul, S. (2016). Dominance and proficiency in early and late bilingualism. In C. Silva-Corvalán, & J. Treffers-Dallers (Eds.), *Measuring Dominance in Bilingualism* (pp. 15–35). Cambridge University Press.
- Simonet, M. (2011). Production of a Catalan-specific vowel contrast by early Spanish-Catalan bilinguals. *Phonetica*, 68(1–2), 88–110. <https://doi.org/10.1159/000328847>
- Tagliamonte, S. A. (2012). *Variationist sociolinguistics: Change, observation, interpretation*. Wiley-Blackwell.
- Trudgill, P. J. (1974). Linguistic change and diffusion: Description and explanation in sociolinguistic dialect geography. *Language in Society*, 3(2), 215–246. <https://doi.org/10.1017/S0047404500004358>
- Trudgill, P. J. (1983). *Sociolinguistics: An introduction to language and society*. Penguin.
- Trudgill, P. J. (1986). *Dialects in contact*. Blackwell.
- Van Coetsem, F. (1988). *Loan phonology and the two transfer types in language contact*. Foris. <https://doi.org/10.1515/9783110884869>
- Wardhaugh, R. (2002). *An introduction to sociolinguistics* (4th ed.). Blackwell.
- Wolfram, W. A. (1969). *A sociolinguistic description of Detroit negro speech*. Center for Applied Linguistics.
- Wolfram, W., & Fasold, R. W. (1974). *The study of social dialects in American English*. Prentice-Hall.