

## Article

# A Phonetic Account of Spanish-English Bilinguals' Divergence with Agreement

Laura Colantoni <sup>1,\*</sup> , Ruth Martínez <sup>1</sup> , Natalia Mazzaro <sup>2</sup>, Ana T. Pérez Leroux <sup>1</sup> and Natalia Rinaldi <sup>1</sup>

<sup>1</sup> Department of Spanish and Portuguese, University of Toronto, Toronto, ON M5S 1K7, Canada; ruth.martinez@mail.utoronto.ca (R.M.); at.perez.leroux@utoronto.ca (A.T.P.L.); natalia.rinaldi@mail.utoronto.ca (N.R.)

<sup>2</sup> Department of Languages and Linguistics, University of Texas at El Paso, El Paso, TX 79968, USA; nmazzaro@utep.edu

\* Correspondence: laura.colantoni@utoronto.ca

Received: 17 September 2020; Accepted: 2 November 2020; Published: 11 November 2020



**Abstract:** Does bilingual language influence in the domain of phonetics impact the morphosyntactic domain? Spanish gender is encoded by word-final, unstressed vowels (/a e o/), which may diphthongize in word-boundary vowel sequences. English neutralizes unstressed final vowels and separates across-word vocalic sequences. The realization of gender vowels as schwa, due to cross-linguistic influence, may remain undetected if not directly analyzed. To explore the potential over-reporting of gender accuracy, we conducted parallel phonetic and morphosyntactic analyses of read and semi-spontaneous speech produced by 11 Monolingual speakers and 13 Early and 13 Late Spanish-English bilinguals. F1 and F2 values were extracted at five points for all word-final unstressed vowels and vowel sequences. All determiner phrases (DPs) from narratives were coded for morphological and contextual parameters. Early bilinguals exhibited clear patterns of vowel centralization and higher rates of hiatuses than the other groups. However, the morphological analysis yielded very few errors. A follow-up integrated analysis revealed that /a and o/ were realized as centralized vowels, particularly with [+Animate] nouns. We propose that bilinguals' schwa-like realizations can be over-interpreted as target Spanish vowels. Such variable vowel realization may be a factor in the vulnerability to attrition in gender marking in Spanish as a heritage language.

**Keywords:** Spanish-English bilinguals; gender; vowels; vowel centralization; vowel sequences

## 1. Introduction

Heritage speakers, i.e., bilinguals who were raised with a home language different from the dominant language of their communities, exhibit various patterns of divergence and cross-linguistic influence. Studies of heritage speakers of Spanish document difficulties with gender agreement and concord for both adults (e.g., Montrul et al. 2008) and children (Gathercole 2002; Montrul and Potowski 2007; Morgan et al. 2013), which may remain at the level of performance (Montrul et al. 2014) or could lead to internal restructuring of the featural system, as proposed by Scontras et al. (2018) and Cuza and Pérez-Tattam (2016). Attrition of the gender system increases across bilingual generations (Martínez-Gibson 2011) and during school years (Cuza and Pérez-Tattam 2016). Given the frequency of the configuration under consideration (a determiner followed by a noun), and the morphological transparency and robustness of the Spanish agreement system, one might speculate that input factors cannot fully account for the vulnerability of this domain beyond lexical errors with low-frequency, gender-opaque nouns. One thing worth keeping in mind is that analyses of bilingual gender consider only some morphosyntactic or lexical factors and leave aside the potential explanatory value of other

domains. We turn to the perception-production interface to explore an alternative explanatory source to current accounts of bilingual vulnerability of Spanish gender.

Spanish agreement is primarily encoded by three vowels (/a e o/) appearing in unstressed word-final position, a context where English neutralizes vocalic contrasts. Word-final vowels followed by other vowels in the word-initial position are another locus of the differences between English and Spanish: English has a tendency towards separating these sequences via insertion of glottal stops or pauses (Davidson and Erker 2014), whereas Spanish diphthongizes or fuses vowels (e.g., Aguilar 2010). Is it possible, then, that heritage bilinguals' realization of Spanish vowels in absolute final position and in sequences introduces what Scontras et al. (2015) dubbed "incipient changes" in the input that could eventually trigger changes in gender marking in subsequent learners' cohorts? Is it also possible that linguists working on bilingual morphosyntax may have overlooked differences in the realization of these vowels when coding bilinguals' speech?

Our overarching goal is not to provide a definitive answer, but to probe preliminary data on these two questions. We conduct analyses of read speech and narratives produced by adult early and late bilinguals (defined in terms of onset of exposure prior to vs. after adolescence, respectively). In our approach, we use phonetic analyses to provide a full characterization of how phonetic variability could potentially impact heritage language acquisition, followed by an exploration of how various lexical and morphosyntactic properties associate with gender realization, and finally, we combine phonetic analyses with accuracy ratings. Thus, we will first phonetically analyze word-final unstressed vowels to compare our results with previous studies (see Section 1.2) that found that early bilinguals centralize vowels when compared to late bilinguals and monolinguals. Second, we will analyze the noun phrases extracted from the narratives to determine the distribution of patterns of morphological realizations in bilinguals to test whether early bilinguals differ from late bilinguals and monolinguals in their rate of errors with gender marking and concord (i.e., gender matching to related morphosyntactic categories). Third, and this is the key contribution of our study, we will integrate our phonetic and morphosyntactic results, arguing that the findings cast some doubt on our capacity of reporting agreement accuracy only based on the researcher's auditory transcriptions. Finally, we will use the insights extracted from this dataset to propose a modular interaction hypothesis, which claims that language contact-related variability in one domain (phonetics-phonology) can have important consequences for the acquisition of other domains (grammar). The testing of this latter hypothesis is a long-term goal of our research team and involves the design of an acquisition study that controls for the perception-production of the target vowels and the comprehension and production of agreement and concord. Before we delve into the analysis of our corpus, we pair previous findings from studies on the acquisition of gender with those of phonetic studies on the acquisition of vowels in Spanish-English bilinguals.

### 1.1. Acquisition of Gender in Spanish-English Bilinguals

Spanish encodes agreement as word-final affixes in the nominal system. Nouns and their dependents (determiners, adjectives, and pronouns) are marked (whether visibly or not) for gender. Gender marking can either have semantic value, as in *niña/niño*, 'girl/boy' or simply be a formal word-marker, as in *pala*, 'shovel' vs. *palo*, 'stick'. Spanish gender is considered transparent as most nouns overtly mark gender by means of the canonical vowel suffix. A few nouns do not directly reflect the gender feature: those ending in consonants, other vowels, or in mismatching vowels.

Monolingual and bilingual children show different patterns of development. Monolingual children learn noun agreement in spontaneous speech by the age of 2;0 (Lopez Ornat 1997; Snyder et al. 2001). Later gender-assignment errors (as identified by the gender of the article; e.g., *\*la mapa*; 'the-FEM map') reflect the lack of lexical knowledge of the gender of opaque nouns and expresses biases towards phonological gender cues (Pérez-Pereira 1991). Errors in concord (matching gender across the noun phrase) are negligible in the preschool years (Castilla and Pérez-Leroux 2010). Bilingual children reach mastery much later (Barreña 1997; Eichler et al. 2013; Larranaga et al. 2012), and early infant speech in bilinguals does not show the harmonic prenominal vowel patterns used by monolingual infants

before they fully acquire determiners (Kuchenbrandt 2005); e.g., a vaca, ‘(the-fem cow)’. Elicited data allow a direct comparison of attainment of noun-adjective agreement in monolingual and bilingual communities. Bedore and Leonard (2001) tested Mexican American five-year-olds in San Diego. These children, described as having minimal access to English, showed noun-adjective agreement accuracy at 76–91%. In contrast, same-age children tested in Mexico by Grinstead et al. (2008) were at ceiling. Morgan et al. (2013) found that five-year-old US bilinguals displayed few, but significantly more, gender substitution errors than their monolingual counterparts. In contrast, Cuza and Pérez-Tattam (2016) presented school-aged children with morphologically opaque nouns. Bilinguals produced many more errors than same-age monolinguals and were delayed in both their knowledge of gender assignment and of concord. Beyond noting a range of patterns of performance across different bilingual populations, some studies link exposure to Spanish to fewer gender errors (Gathercole 2002). For older children, Montrul and Potowski (2007) found differences between monolinguals and bilinguals, and between simultaneous and sequential bilinguals.

More recent studies (Goebel-Mahrle and Shin 2020), however, failed to find differences between monolingual and bilingual heritage speakers, independently of their age, which ranged between 5 and 11. This was also the case in the word-repetition task included in Montrul et al. (2014) study. Although adult Spanish heritage speakers and L2 learners showed differences from the monolingual comparison group in the other two tasks in the study (a gender monitoring task and a grammaticality judgement task), they did not differ from monolinguals in the word-repetition task. While acknowledging that bilinguals are not a uniform group, we note that the results of the word-repetition task might not necessarily indicate a higher accuracy in production by bilinguals, but in the limitations of auditory transcriptions instead. From a performance point of view, gender is often highly predictable in context. If bilinguals produce a schwa-like vowel rather than the underlying vowel, a transcriber might analyze the item as correct, independently of the quality of the target vowel. A schwa-like vowel insertion allows to maintain the syllable structure of the target word, and introduces acoustic ambiguity. In a predictable context, the transcriber perceives a vowel that cannot be clearly interpreted as a non-target vowel (e.g., *caperucita rojo*, ‘Little Red Riding Hood<sub>FEM</sub> red<sub>MASC</sub>’) and transcribes the vowel as [a], in what can be interpreted as a case of *in dubio pro reo*. Why do we think this is the case? Because the literature that we will review below consistently shows that Spanish-English bilinguals exhibit clear patterns of centralization of their vocalic space. We strongly believe that pairing up the phonetic literature on the acquisition of vowels by Spanish-English bilinguals with the findings from the literature on the acquisition of gender by the same group of bilinguals should at least make us wonder if the findings are compatible.

### 1.2. Acquisition of Vowels in Spanish-English Bilinguals

Spanish vowels (/a e i o u/) vary little in quality and remain contrastive in stressed and unstressed positions (Hualde 2014). Although the literature reports variation in duration (e.g., Delattre 1965) and, to a lesser extent, in quality (Romanelli et al. 2018), Spanish does not show centralization to the extent of English (Navarro Toma’s 1970, p. 43). Spanish vowels are learned early, typically by the age of 2;0 (Goldstein and Pollock 2000; Schnitzer and Krasinski 1994). This contrasts with English, where infants begin with a centralized vowel space that expands over time (Gilbert et al. 1997; Kent and Murray 1982; Rvachew et al. 1996). Vowel space differentiation is reported for eighteen-month-olds (Rvachew et al. 2006), but the full English vocalic inventory is mastered later, by 3;0 (Stoel-Gammon and Sosa 2008; Stoel-Gammon and Pollock 2009). This is to be expected, given that English has a large vocalic inventory in the stressed position (Ladefoged 2001). In the unstressed position, the only frequent vowel is schwa [ə] (Rogers 2000). Thus, the unstressed vowel inventory is smaller in English than in Spanish (Hualde 2014).

Studies on the acquisition of Spanish vowels by bilinguals (Ronquest and Rao 2018) report consistent effects of cross-linguistic influence. Bilingual children centralize unstressed vowels

(Gildersleeve-Neumann et al. 2009; Menke 2010),<sup>1</sup> although it is debated whether centralization happens before or after children enter the school system (Gildersleeve-Neumann et al. 2009). Differences in vowel realizations persist into adulthood (Rogers 2012; Ronquest 2016; Willis 2005). Cross-linguistic effects are also reported in perception. Studies of English learners of Spanish suggest late bilinguals tend to confuse high-front with mid-front vowels (/i/ with /e/) and back-mid with low vowels (/o/ with /a/) (Morrison 2003; Morrison 2006). Unlike monolinguals, bilinguals rely on frequency cues rather than on duration (Fox et al. 1994). One perception study on adult heritage Spanish speakers (Mazzaro et al. 2016) found differences, but only in unstressed positions.

These results, however, refer to single vowels. A factor not yet considered is how phonological processes affect word-final vowels, the cross-linguistic differences in this domain, and the potential cross-linguistic interactions. Spanish tends to fuse vowels across words. When both vowels are unstressed, as in *como alfajores* ('I eat cookies'), the highest vowel in the sequence is frequently reduced, ranging from gliding to full deletion. Reduction processes apply to high and non-high vowels equally (Hualde et al. 2008; Vokic and Guitart 2009) across Spanish dialects (Alba 2006; Hutchinson 1974), but reduction is less frequent when one of the vowels is stressed (Colantoni and Hualde 2016; Hualde 2014). Because agreement vowels are in the unstressed word-final position, they are likely modified in running speech. Contrastingly, vowel reduction across word boundaries is rare in English, given that speakers frequently insert glottal stops to separate across-word vocalic sequences (Davidson and Erker 2014). Thus, whereas Spanish prefers diphthongization or deletion, English realizes these sequences as hiatuses.

Crucially, for coarticulation to occur, words must belong to the same intonational phrase or the same prosodic unit. Because nouns and adjectives are prosodified together in Spanish (D'Imperio et al. 2005; Frota et al. 2007), we assume that agreement vowels are often coarticulated in the input. An important question is whether coarticulation affects bilingual patterns of acquisition of suffixal morphology. It has been shown that word-edges play an important role in word recognition (Shoemaker and Rast 2013) and that coarticulation of consonants has a negative effect on lexical retrieval, as speakers struggle to compensate for coarticulated sounds (Mohaghegh 2016). We can thus expect a similar effect for vowels.

### 1.3. Phonetics and Morphosyntax

Research on the bilingual acquisition of Spanish morphology has explored multiple factors, such as frequency of use, nature of the input, and type of target form, but has not examined the phonetic properties of agreement. This was previously pointed out by Silva-Corvalán (2014). Phonetic factors, however, are known to be important predictors of functional morphology in the child acquisition literature. Prosody is a common explanation of omission of functional elements, such as articles and clitics (the Prosodic Bootstrapping hypothesis by Lleo' and Demuth 1999; Guasti et al. 2008; Mateu 2015). A robust literature reveals that phonology is key in morphological development. For example, accuracy in production of the English plural -s and third person -s depends on syllable (coda) structure (Ettlinger and Zapf 2011; Song et al. 2009; see also Bernhardt and Stemberger 1998 for an overview). Culberston et al. (2019) show that children use phonological rather than semantic categories when acquiring gender. Phonetic reduction appears to have an impact on the overall course of development, affecting both comprehension and production. Miller and Schmitt (2010) show that the variable realization of the final /s/ in Spanish impacts the acquisition of both plurals and tense agreement (2Sg present). Children growing up in varieties where /s/ is maintained (i.e., not weakened or deleted) acquire these markers earlier, both in comprehension and production, than children growing up in varieties where /s/ is frequently aspirated and deleted. Thus, given evidence that phonetic characteristics of the input affect the acquisition of number (English, Spanish) and person (English),

<sup>1</sup> Kehoe and Lleó (2017) report similar results for German-Spanish bilinguals.

we hypothesize that phonetic variability induced by cross-linguistic influence in the phonetic domain will impact the acquisition of gender in Spanish-English bilinguals.

#### 1.4. Research Questions and Hypotheses

RQ1: What is the phonetic realization of word-final unstressed vowels and across-word vowel + vowel sequences in the three groups?

H1: We hypothesize that early bilinguals (EB) will show a higher rate of vowel overlap (single vowels) than the other two groups. We also hypothesize that monolinguals (M) and late bilinguals (LB) will tend to fuse vowels across words while EBs will tend to separate them.

RQ2: Are there more errors in gender agreement and concord in EBs than in the other groups?

H2: EBs will show a larger proportion of errors than the other two groups.

RQ3: Is vowel centralization being reported as accurate gender marking?

H3: EBs will show a higher rate of centralization when compared with the other two groups, and cases of vowel centralization will be labelled as accurate gender.

RQ4: Do contextual factors (i.e., predictability of the gender of the noun) predict vowel centralization?

H4: Highly predictable DPs are less likely to be fully specified phonetically as they offer redundant information.

## 2. Materials and Methods

### 2.1. Participants

A total of thirty-seven participants ( $N = 37$ ) took part in the study: 13 Early Bilinguals, 13 Late bilinguals, whose first language was Spanish, and 11 monolingual speakers serving as the comparison group. All the bilingual participants were residents of El Paso, US who were attending different classes at the University of Texas at El Paso. Participants in the (functionally) monolingual group were Spanish speakers with minimal exposure to and ability in English and were residents of Ciudad Juárez, Mexico. They were either recruited from a beginner-level English for Speakers of Other Languages (ESOL) course, or contacted through social networks.

Following previous research (Montrul 2011; Silva-Corvalán 2014), the EBs were second- or first-generation immigrants who acquired Spanish during childhood at home or in other natural contexts where a majority language (English) was spoken. They were either born and raised in the US or immigrated permanently to the US at or before the age of 12. The LBs were of first generation Mexican background who arrived in the US after the age of 13 with fully developed L1 grammar. Participants completed an adult language background questionnaire, which elicited information on place of birth, primary language of schooling, patterns of language use, etc. This questionnaire also elicited a self-proficiency judgment in both English and Spanish in the four linguistic skills via a Likert scale, ranging from basic/limited (1) to excellent/native (4). In addition to the self-proficiency measure, participants completed an independent proficiency task, adapted from the *Diploma de Español como Lengua Extranjera* (DELE) (Cuza et al. 2013). Table 1 summarizes Age, Age of Arrival to an English context (AOA), and Length of Residence in an English context (LOR) for all participant groups.

EBs scored an average of 40 in the DELE test, while LBs scored 45. Previous research using this methodology (Cuza et al. 2013; Montrul et al. 2003) considered participants who scored between 40 and 50 points (out of 50) to be ‘advanced’ learners, those with scores of 30 to 39 were considered to be ‘intermediate’ learners, and those with scores between 0 and 29 were considered to be ‘beginner’



learners. In other words, while the proficiency score of the EB group is a bit lower than other groups, all groups have a high overall level of proficiency in Spanish. The EB group included participants born and raised in the US and those who came to the US before the age of six (AOA = 2.5). Their self-proficiency rating in English was near native (3.73/4), while, in Spanish, it was good/fluent (2.9/4). Regarding their patterns of language use, most of the participants (62%) reported using both English and Spanish at home, but they used mainly English at work (75%) and in social situations (46%). Six participants felt more comfortable in English (46%) and six participants (46%) felt equally comfortable in both English and Spanish (46%). Only one participant (8%) selected Spanish as the most comfortable language.

**Table 1.** Participants' demographic information.

	Early Bilinguals (EB) (n = 13; Female = 7)			Late Bilinguals (LB) (n = 13; Female = 9)			Monolinguals (M) (n = 11; Female = 10)		
Mean age at testing (SD)	23 (2.64)			38 (13.46)			25 (11.13)		
Mean AOA (SD)	2.5 (3.6)			23 (9.37)			20 (4.23)		
Mean LOR (SD)	21 (2.95)			16 (9.6)			08 (0.8)		
Self-Proficiency	English = 3.73/4 Spanish = 2.9/4			English = 2.9/4 Spanish = 3.9/4			English = 1.43/4 Spanish = 4/4		
DELE score	40/50			45/50			45/50		
Language use	SPAN	ENG	BOTH	SPAN	ENG	BOTH	SPAN	ENG	BOTH
Home	15%	23%	62%	77%	8%	15%	100%	0%	0%
School	8%	77%	15%	42%	17%	41%	90%	0%	10%
Work	9%	64%	27%	10%	40%	50%	100%	0%	0%
Social situations	23%	46%	31%	58%	0%	42%	100%	0%	10%
Most comfortable in	SPAN	ENG	BOTH	SPAN	ENG	BOTH	SPAN	ENG	BOTH
	8%	46%	46%	77%	0%	23%	100%	0%	0%

Late bilinguals ( $n = 13$ ) included first generation immigrants from Mexico (mean age at testing = 38; mean AOA = 23; mean LOR = 16). Their self-proficiency rating in Spanish was almost native (3.9/4), and in English it was good/fluent (2.9/4). The proficiency score in the DELE test was 45/50 (advanced proficiency). As for language use, the majority reported speaking more Spanish at home (77%) and in social situations (58%). At school, five participants ( $N = 5$ ) used both English and Spanish and another five used only Spanish (42%). At work, half of the participants used only English, and 40% used both English and Spanish. When asked which language they felt most comfortable in, the majority (77%) indicated Spanish.

The comparison group consisted of seven recent arrivals in El Paso, Texas, and four residents of Ciudad Juárez, Mexico (mean age at testing = 25; mean AOA = 20; mean LOR = 9 months). Although their AOA of English is earlier than LB (20 vs. 23), these speakers had learned English for a shorter period of time, specifically an average of 9 months. Most of the participants reported speaking more Spanish at home, school, work, and in social situations. They also reported feeling most comfortable speaking Spanish.

All subjects gave their informed consent for inclusion before they participated in the study. The study was conducted in accordance with the Declaration of Helsinki, and the protocol was approved by the University of Texas at El Paso IRB.

## 2.2. Materials and Tasks

To test our hypothesis, we analyzed reading data from "The North Wind and the Sun" and the narrative of the folk tale "Little Red Riding Hood". To elicit the narrative, participants were shown wordless pictures of a children's book based on Perrault's version of the tale. First, participants were shown the pictures as a refresher. When they felt ready, participants recorded the narrative using the images as guidance. All the sessions were conducted in a sound-treated room. Informants were recorded directly onto a laptop computer using Audacity 2.1.2 (available on-line: <http://audacityteam.org/>) and a Blue Snowball USB microphone. The speech was sampled at 44.1K, and word-final vowels and vocalic sequences (/a e o ae ea oa/) were analyzed with PRAAT (Boersma and Weenink 2001).

### 2.3. Analysis

The data obtained was subject to two analyses. To analyze the vowels phonetically, we extracted all tokens of /a e o/ in unstressed word-final position as well as all the unstressed across-word vowel sequences (the first vowel in the sequence was one of the three target vowels in the study) using PRAAT. We marked the onset of each vowel or vowel sequence at the beginning of the F1 increase and the offset at the drop in intensity (pre-pausal vowels) or at the beginning of the F1 rise for cases in which the word-final vowel was immediately followed by a word beginning with a consonant. All measurements were taken at zero-crossings. From the reading, we extracted the vowels from nouns, adjectives, and verbs (single vowels:  $N = 1045$ ; sequences:  $N = 287$ ) whereas, from the narrative, only vowels that were part of the noun phrase were extracted (single vowels:  $N = 2954$ ; sequences:  $N = 1109$ ).<sup>2</sup> Although formant values were automatically extracted at five points, we will report values at mid-point for single vowels (to minimize the effects of coarticulation with surrounding consonants) and values at five points for sequences. Formant values were subsequently checked manually to inspect values that fell outside the ranges reported in previous studies for Spanish and Spanish-English bilinguals. Data were normalized using the Lobanov (1971) method, as adapted by Nearey (1977) and Adank et al. (2004). Single-vowel results were submitted to two complementary analyses aiming at computing the degree of overlap in vocalic spaces in each group. The first analysis computes the Bhattacharyya affinity scores, following Johnson (2015) and Strelluf (2016). The Bhattacharyya's affinity measures the degree of overlap between two Gaussian distributions (Mak Brian 1996) and is considered to be less sensitive than other approaches when there are imbalances in the sample (Strelluf 2016), which is the case in both tasks but particularly in the narrative. A score of 0 indicates that there is no overlap between vocalic spaces, whereas a score of 1 signals complete overlap. To calculate the Bhattacharyya's affinity score we used the "kernel overlap" function of the {adehabitatHR} R library (Calenge 2006). The second analysis, the computation of convex hulls, as implemented by Haynes and Taylor (2014), allowed us to quantify the percentage of overlap between vowel pairs by calculating the smallest convex shape (polyhedron) that would fit all the given datapoints (Haynes and Taylor 2014, p. 885).<sup>3</sup> Results are reported as percentages and were calculated with the package {phonR} (McCloy 2016).

Vocalic sequences were labeled and then transcribed by one of the authors and then verified by another author. Using the acoustic information available in the spectrogram, we distinguished sequences that were realized together (i.e., no pause or prosodic brake, such as pitch reset between the vowels) from those that were separated by a pause or a glottal stop. We labelled the former as 'diphthong' and the latter as 'hiatus'. This allowed us to clearly separate the sequences that cannot be coarticulated (those labeled as 'hiatus') from those that could (those labeled as 'diphthong'). Thus, the sequences labeled as 'diphthong' may include hiatus realizations, i.e., sequences in which each vowel has frequential and durational values that are similar to those of single vowels (e.g., Aguilar 1999; Borzone de Manrique 1976; Colantoni and Limanni 2010).<sup>4</sup> To determine whether there were frequential differences, sequences that were labeled as diphthongs were acoustically analyzed for their formant trajectories. For the analysis of F1 and F2 trajectories within the unstressed sequences /ae oe ea/,<sup>5</sup> we ran Smoothing Splines ANOVAs (SSANOVAs) using the package {gss} (Gu 2014) to test for statistical differences across the three groups of speakers (EB, LB, M) at five intervals within each sequence. SSANOVAs create smoothing splines for each group by connecting mean data points

<sup>2</sup> Although no tokens had to be discarded from the reading task, we excluded 72 single vowel tokens from the narrative which were produced with creaky voice.

<sup>3</sup> As explained by Haynes and Taylor (2014), these are measurements of overlap and not of statistical significance. As in Haynes and Taylor's study, we are not interested in statistical differences but in the extent to which vocalic spaces overlap, because a greater degree of overlap suggests that the vowels are articulated similarly.

<sup>4</sup> Hiatus also differ from diphthongs in the duration and trajectory of formant transitions (e.g., Aguilar 1999).

<sup>5</sup> Only these sequences were analyzed because (i) the first vowel was unstressed and (ii) they appeared in the reading task and in the narrative.

through time, as well as 95% Bayesian confidence intervals represented by dotted curves above and below the splines. At the time-points where confidence interval curves do not intersect, two splines are considered significantly different.

To conduct the morpho-syntactic analysis, all noun phrases (NPs) were extracted, representing a total of 2445 analyzed tokens.<sup>6</sup> Determiner phrases (DPs) were coded for configuration (whether it contained a noun, determiner or quantifier, or modifiers, as in (1)), realization of gender and number agreement in each relevant category (nouns, determiners, and adjectives), and overall concord patterns (match/mismatch between constituents as well as target realization for the head noun, as shown in (2)). All nouns were further analyzed for various semantic parameters to explore potential association between form and meaning. These included semantic type of noun in terms of categories pertaining to concreteness and animacy (3) and to individuation and countability (4). Additional coding included whether the noun's initial segment was stressed [a], which, in Spanish singular feminine nouns, leads to an exceptional use of the masculine article (5), as in *el águila* (f) 'the eagle'.<sup>7</sup> Overtly marked gender is indicated in the glosses.

(1) DP configuration

<i>Noun:</i>	lobo	'wolf'
<i>Adjective + Noun</i>	gran lobo	'big wolf-MAS'
<i>Determiner + Noun:</i>	el lobo	'the-MAS wolf-MAS'
<i>Determiner + Covert Noun + Adjective:</i>	el grande	'the-MAS big'
<i>Determiner + Adjective + Noun:</i>	el gran lobo	'the-MAS big wolf-MAS'
<i>Determiner + Noun + Adjective:</i>	el lobo grande	'the-MAS wolf-MAS big'

(2) Agreement: match for agreement features (number or gender) with determiners and/or adjectives, and with target gender of the noun

<i>Correct:</i>	las casas bonitas	'the-FEM pretty-FEM houses-FEM'
<i>Incorrect:</i>	la casa bonito	'the-FEM pretty-MAS house-FEM'

(3) Noun semantic type: semantic features of the noun under analysis

<i>Abstract:</i>	paz	'peace'
<i>Concrete:</i>	coche	'car'
<i>Animate:</i>	gato	'cat'
<i>Human:</i>	chica	'girl'
<i>Event:</i>	fiesta	'party'

(4) Individuation

<i>Mass:</i>	arena	'sand'
<i>Individual:</i>	coche	'car'
<i>Collective:</i>	equipo	'team'
<i>Ambiguous (used as individual):</i>	la policía	'female police officer'
<i>Ambiguous (used as collective):</i>	la policía	'police'
<i>Both (mass or individual):</i>	fruta	'fruit'

All nouns were also coded for morphology and for their morphological relation to other noun lexemes. First, we isolated the final suffix or segment. The first goal was to determine whether gender was visibly expressed or not; that is, whether the noun contained the transparent word markers -a

<sup>6</sup> Frozen expressions such as *Fin*, *Colorín Colorado* ('the end') and *lugar* ('place') in the prepositional locution *en lugar de* ('in lieu of') were excluded from the analysis.

<sup>7</sup> The morphosyntactic coding was conducted by one of the authors and then was verified by another author. As concerns the error count, this was first checked by one investigator, then re-checked by a second author, and discrepancies (N = 2) were solved by a third author.



(f) and -o (m), a different transparent suffix such as *-ción*, which is uniformly feminine, or whether it contained formally opaque final segments or suffixes (5). We then considered whether the noun entered a gender alternation or not, as in (6), and if so, what was the lexical relationship to the other entry in the alternation (7).

(5) Gender visibility

[nouns with formally visible gender]

<i>word marker:</i>	niño (m)	'boy'
	niña (f)	'girl'
<i>transparent suffix:</i>	educación (f)	'education'
	cazador (m)	'hunter'
[nouns with no visible gender marking]		
<i>-e:</i>	padre (m)	'father'
<i>-i / -u:</i>	espíritu (m)	'spirit'
<i>consonant:</i>	amor (m)	'love'
<i>inverted (-o for fem. /-a for masc.):</i>	mano (f), foto (m)	'hand, photo'

(6) Gender alternation: whether or not there exists an opposite word to pair with the noun in terms of gender and the nature of the lexical relationship between the alternants

<i>Transparent:</i>	gato (m)/gata (f)	'male cat, female cat'
<i>Unmarked masculine:</i>	jefe/jefa (m)/(f)	'male boss, female boss'
<i>Lexical root gender:</i>	hombre (m), mujer (f)	'man, woman'
<i>Derivational:</i>	cerezo/cereza (m)/(f)	'cherry tree, cherry'
<i>Free variation (alternating forms with minimal or no semantic change):</i>		
	canasta (f)/canasto(m)	'basket'
<i>Unrelated:</i>	plato (m)/plata (f)	'plate, silver'
<i>Invisible gender (semantically alternating but marked only in article):</i>		
	estudiante (f)/(m)	'male or female student'
<i>Unique (only one gender per root):</i>	carro (cf *carra)	'car'
<i>Episone (one grammatical gendered semantically unspecified gender):</i>		
	víctima (f)	'male or female victim'

We finally considered both structural and referential context to assess whether the gender of a given DP was predictable. Predictability was categorical (y/n), and we annotated the source of predictability, either by syntactic (informative article) or by semantic/contextual means (known/previously mentioned). The purpose of this classification was to explore whether predictability of the gender of a DP predicts vowel underspecification.

Our final analysis combined the phonetic and the morphosyntactic results. The F1 and F2 values obtained for single word-final vowels in the narrative were Bark-transformed and combined with the morphosyntactic analysis.<sup>8</sup> In order to visualize the results, we created a new independent variable that was the result of subtracting the F1 (Bark) to the F2 (Bark). This allowed us to compare the degree of vowel centralization for each participant (the smaller the number, the more centralized the vowel). Then, we plotted F2 and F1 against the predictability of each noun and displayed the results by vowel organized by group. This visualization allowed us to explore the hypothesis of whether nouns that entered into predictable alternations showed a smaller degree of centralization than nouns whose gender is not predictable. To determine whether predictability played a role in vowel centralization, we ran linear mixed effects models with F2-F1 (in Bark) as the dependent variable, Vowel (/a e o/), Predictability (yes/no) and Group (Monolinguals, EB, LB) as independent variables, and Participant as a random factor. All statistics were calculated with [R Studio Team \(2020\)](#).

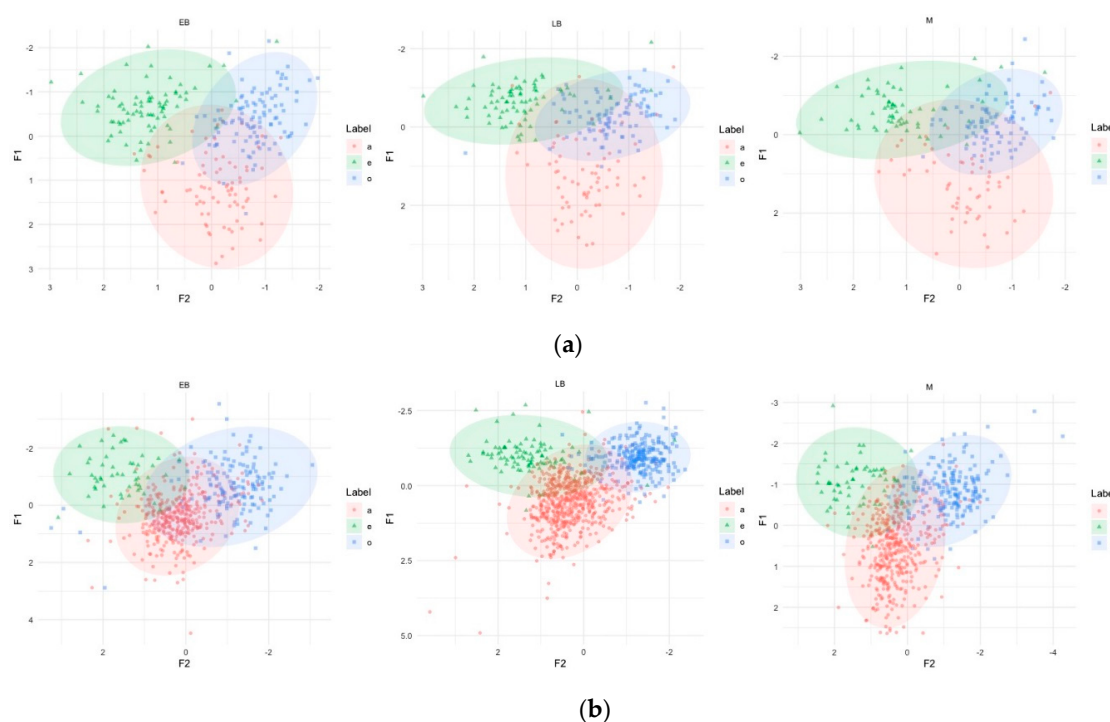
<sup>8</sup> The combined analysis only included data from the narrative because we are interested in comparing vowel quality and gender accuracy. The reading task, instead, is not clearly reflecting participants' grammatical knowledge.

### 3. Results

We begin this section by summarizing the results of the phonetic analysis, which includes the characterization of single vowels and vowel sequences. Section 3.2 presents the morphosyntactic analysis and the final section combines the phonetic results obtained for single vowels with morphological predictability to explore whether it may be possible that vowel centralization is being undetected when we report accuracy in gender agreement.

#### 3.1. Phonetic Analysis

Figure 1a displays the F1–F2 normalized values obtained for single vowels in the reading task, whereas Figure 1b shows the results for the narrative. In both graphs, we observe some overlap, particularly between the vocalic spaces for /a/ and /o/ for all three groups. In the reading task, although bilinguals show slightly more overlap than monolinguals, the patterns are rather similar across groups. In Figure 1b, instead, both bilingual groups display a smaller vowel space for /a/, and EBs clearly show a greater degree of overlap between /a/ and /o/ realizations. This is important because these are the two vowels that are frequently used to encode gender.



**Figure 1.** Formant charts for /a e o/ in all groups (Early Bilinguals (EB), Late Bilinguals (LB) and Monolinguals (M)): (a) reading task; (b) narrative.

Table 2 summarizes two measurements of overlap; the Bhattacharyya’s affinity scores and the percentage of overlap, calculated using convex hulls for the data obtained from the reading passage, whereas Table 3 displays the results of the narrative. As indicated in Section 2.3, both are complementary measurements to quantify overlap. Whereas the affinity score quantifies how similar the vocalic spaces are, the convex hulls quantify the degree of overlap. In both cases, the larger the number, the greater the overlap. As mentioned, the goal of these measurements is to quantify overlap, not to test statistical significance among groups.

These results show, first, a task effect and, second, a vowel-pair effect. The degree of overlap across pairs of vowels (Table 2) is lower in the reading task than in the narrative, and this is particularly evident in the values calculated using convex hulls. More careful articulation and, thus, less overlap is generally expected in a reading task. The degree of overlap, instead, increases in the narrative

at different rates across groups (Table 3). LBs and particularly EBs double and sometimes triple (e.g., [a]–[o]) the degree of overlap in vocalic spaces when speaking spontaneously (see overlap %), which suggests that the distinction between some pairs of vowels may be weakening. This is clearly the case with the [a]–[o] pair, and, to a lesser extent, with the [a–e] pair.<sup>9</sup> In turn, this tells us that the most frequent vowel pairs that mark gender are being produced with the same quality a third of the time by EBs.

**Table 2.** Affinity scores and proportion overlap (convex hulls) for /a e o/ in the reading task. Results displayed by group (Early Bilinguals (EB), Late Bilinguals (LB) and Monolinguals (M)).

Group	Affinity (0–1)			(Convex Hulls) Overlap (%)		
	a–e	a–o	e–o	a–e	a–o	e–o
EB	0.23	0.49	0.21	13.8	10.2	11.8
LB	0.35	0.48	0.28	15.4	11.9	11.5
M	0.31	0.40	0.36	15.8	10.5	9.8

**Table 3.** Affinity scores and proportion overlap (convex hulls) for /a e o/ in the narrative. Results displayed by group (Early Bilinguals (EB), Late Bilinguals (LB) and Monolinguals (M)).

Group	Affinity (0–1)			(Convex Hulls) Overlap (%)		
	a–e	a–o	e–o	a–e	a–o	e–o
EB	0.36	0.52	0.17	26	32.1	21.3
LB	0.35	0.24	0.17	27.4	25.1	15
M	0.36	0.41	0.16	17.5	18.9	17.8

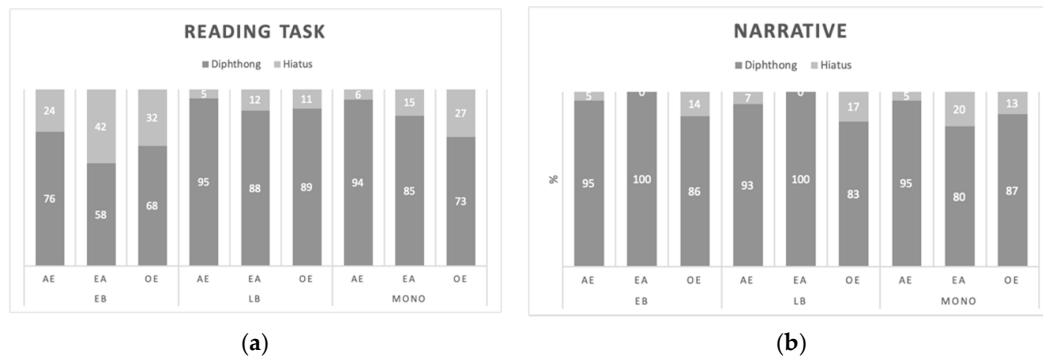
We turn to the realization of sequences, which were analyzed independently from single vowels because the formant quality and trajectory largely depends on how they are realized. If realized as hiatuses, formant values should be similar to those obtained for single vowels, whereas values in diphthongs should differ from the values of the corresponding single vowels (e.g., [Borzone de Manrique 1976](#); [Aguilar 1999](#)). Thus, before conducting the acoustic analysis, it is important to determine the proportion of sequences produced together from those realized in different syllables (Figure 2). A preference for hiatuses will be interpreted as indicative of cross-linguistic influence, since across-word vowels in Spanish tend to be pronounced together (e.g., [Aguilar 2010](#)), which is not the case in English ([Davidson and Erker 2014](#)). The fact that vowels are syllabified differently across words in Spanish has also important implications for L1 acquisition because, as mentioned, the vowel quality in diphthongs differed from the vowel quality in singleton vowels. Thus, a higher probability of diphthongs is also a higher probability of noise in the signal for bilinguals and, thus, for difficulty in determining the quality of the vowel that marks gender. Results in Figure 2a show that EBs, in the reading task, had a larger proportion of hiatuses than the other groups, whereas all groups behaved similarly in the narrative (Figure 2b).

To determine if there were differences in the realization of sequences labeled as diphthongs, we compared the formant trajectories across groups.<sup>10</sup> First, we analyzed the sequences produced by female speakers in the reading task. Male speakers’ productions were not analyzed given the small number of tokens obtained for this task (see Methods) and the even smaller number of tokens in which vowels were not separated by a pause or glottal stops. In terms of F1 trajectories, no statistical differences were observed across the three groups (Appendix A, Figure A1). In terms of F2 trajectories, the SSANOVAs revealed significant differences for all sequences, where the EB group realized (i) /ae/

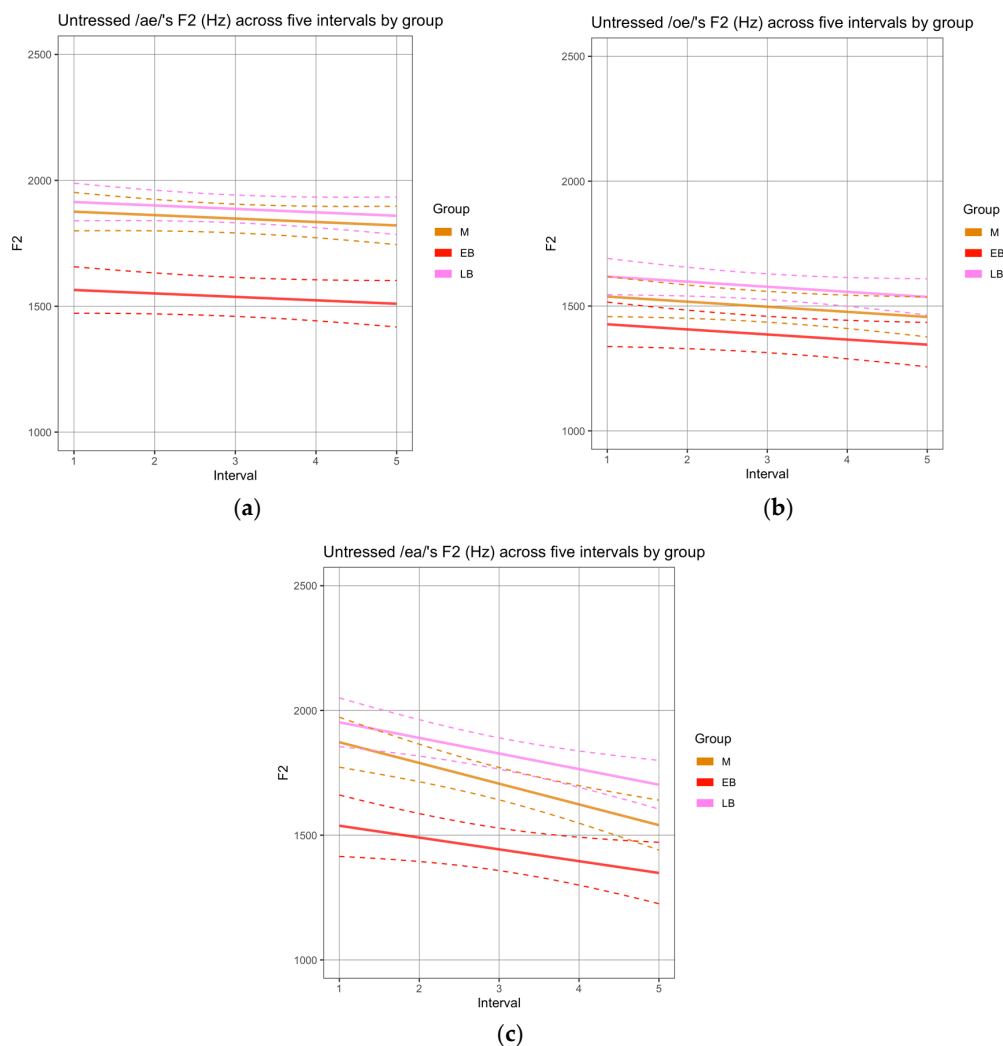
<sup>9</sup> Results regarding the vowel [e] in the narrative should be interpreted with caution, because this vowel represents only 10% of the total tokens.

<sup>10</sup> Note that only sequences that were realized as diphthongs in all groups were analyzed.

with significantly lower values than M and LB (Figure 3a), (ii) /oe/ with significantly lower values than LB (Figure 3b), and (iii) /ea/ with significantly lower values than M and LB (Figure 3c). The F2 trajectories realized by the M and LB groups did not differ significantly.

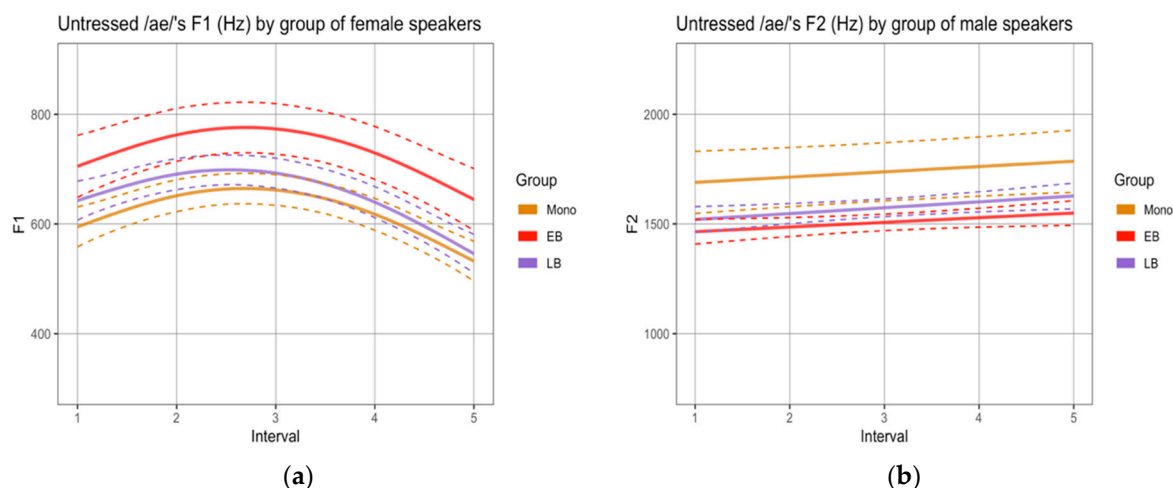


**Figure 2.** Percentage of diphthongs and hiatuses by type of vowel sequence and group (horizontal axis): (a) Reading task; (b) narrative.



**Figure 3.** Smoothing Splines ANOVAs (SSANOVAs) for non-normalized (female) F2 trajectories across five intervals produced by each group of speakers (M, EB, LB) in the reading task. (a) Sequence /ae/; (b) sequence /oe/; (c) sequence /ea/.

Second, and in order to be able to compare results across tasks, we analyzed the sequences produced by female and male speakers in the narrative separately. SSANOVAs by gender using non-normalized values revealed two statistical differences, both within the sequence /ae/: (i) the female EB group realized F1 trajectories with significantly higher values than the female M group (Figure 4a); and (ii) the male EB group realized F2 trajectories with significantly lower values than the male M group (Figure 4b).<sup>11</sup>



**Figure 4.** SSANOVAs for non-normalized formant trajectories across five intervals within the sequence /ae/ produced by each group of speakers (M, EB, LB) in the narrative task: (a) F1 trajectories: female speakers; (b) F2 trajectories (male speakers).

Taken together, these results suggest that the EB group behaves differently from the M and LB groups, often centralizing F2 values in unstressed vowel sequences.

### 3.2. Morphosyntactic Analysis

The grammatical analysis showed high accuracy counts. From a total of 2445 DP tokens extracted for analysis, we identified six clear gender errors (see (7)–(12) below), as well as another handful of errors with grammatical number, mostly singular in lieu of plural, which will not be discussed here. Erroneous marking is indicated in bold.

- (7) la abuelita y la caperucita rojo ... (Speaker UT052, EB)  
The-FEM grandmother-FEM and the-FEM little-riding-hood- FEM red-MASC
- (8) Y la niña la caperucita rojo ... (Speaker UT052, EB)  
And the-FEM girl-FEM the-FEM little-riding-hood-FEM red-MASC
- (9) con **ello** y a a le hablaba y ... (Speaker UT054, EB)  
With **him** and ah ah her-DAT spoke and [intended reference to “with **her**”]
- (10) bueno su abuelo la había ... (Speaker UT107, M)  
Well her grandparent-MASC her-DAT had
- (11) todas **las** alimentos ... (Speaker UT086, LB)  
All-FEM the-FEM food-MASC
- (12) tanto zozobra ... (Speaker UT002, LB)  
So-much-MASC unstability-FEM

We observed two additional, more ambiguous instances where a demonstrative could function as a filler or a stranded phrase or a demonstrative. In (13), below, we note that Speaker UT030 was a

<sup>11</sup> Non-significant results for female and male speakers are included in Figures A2 and A3, respectively.



frequent user of the *este* filler. Nonetheless, there was no prosodic break separating the noun from the preceding demonstrative, so an analysis of these tokens as instances of determiner-noun disagreement cannot be ruled out.

(13) *este caperucita* está entrando ... (Speaker UT030, M)

This-MASC little-riding-hood-FEM is entering

(14) *y eso* cosas como *eso* ... (Speaker UT054, EB)

And this-MASC things-FEM like this-MASC

This high accuracy rate is appropriate, given the high proficiency status of these speakers, many of which produced elaborate, lexically rich and syntactically complex narratives that were, for the most, seemingly error free and without English intrusion. Nonetheless, the formal and contextual properties of gender expression in the DP deserve further examination. We ask two questions: How robustly do these DPs manifest gender agreement explicitly? And given our phonetic results, how accurate will a coder or transcriber be at detecting that a given gender vowel has been centralized. If a gender form is highly predictable (given lexical retrieval of a contextually or syntactically predictable entry), a coder might be likely to perceive a centralized vowel as the target. Given what is known about speech perception, we can expect that any skilled listener will predict a centralized vowel (*caperucit@*) to be an /a/ and ignore the presence of the schwa. Even if we think we are paying attention, the evidence suggests otherwise. Listeners are generally known to perceive elements that are not present in the speech chain in order to repair phonotactically illicit sequences (e.g., [Calabrese 2012](#); [Durvasula and Kahng 2015a](#); [Durvasula and Kahng 2015b](#); [Hawkins 2010](#); [Repp 1992](#)); so, we might expect them to easily attribute features to underspecified central vowels.

To further explore these patterns of accuracy, we consider the distribution of lexical and morphological types of noun, determiner, and adjective forms. Nouns were fairly evenly divided between those that entered into a gender alternation (52%, 1284 tokens) and those that did not; and, for this particular story, feminine nouns were almost twice as frequent as masculine nouns (1575 vs. 885 tokens). For most nouns, gender marking is an arbitrary classification with no semantic import. The main exceptions are nouns referring to human entities, some but not all animal classes (cf. *gallina* vs. *avispa*), and a handful of narrow subclasses of systematic alternations (*cerezo/cereza*). Nouns with human referents (*abuela/madre/caperucita/niña/cazador*) made up 42% of the data; animals (mostly *lobo*, ‘wolf’, with an occasional reference to cats in *Caperucita*’s house) added another 12% to the count of semantically transparent gender alternations.

From a lexical perspective, over 80% of the noun tokens (2028 tokens) in the narratives had explicit gender marking (i.e., transparent *-o/-a* word-markings). Other nouns had either opaque word-final morphology (ending in *-e* or consonant) or ended in a transparent suffix (*-ción*, etc.). A potential loss of the gender system is revealed either as (i) a switch in the word marker vowel (i.e., saying *caperucito* or *abuelo* in lieu of *abuela*), or (ii) a mismatch between the noun and the determiner (*las alimentos*). A determiner-noun mismatch can potentially indicate loss of concord or agreement; most likely, it may just represent a lexical knowledge gap: the speaker may have misclassified the gender class to which a noun belongs. A study of gender agreement in French-English bilingual children by [Nicoladis and Marchak \(2011\)](#) supports this claim. Their data showed less attrition in concord (i.e., agreement between determiner-adjective) than in gender assignment (i.e., what determiner was associated to a noun). Bilingual children were significantly less accurate with gender assignment beyond their differences in lexical scores. However, when these authors strictly considered concord, there were no statistical differences between monolingual and bilingual children.

Determiners as a category were only partially informative. Bare nouns were common (26%, 629 tokens), and some determiners such as possessives (*mi casa*) and quantifiers (*dos casas*) were uninformative for gender (14%, 347 tokens), so that about 40% of the analyzed tokens had no gender identification in the determiner. The determiner forms that inflect for gender (*definites*, *indefinites*, *demonstratives*) can be divided into those where gender identifiability depends only on perceptibility of

the gender vowel, as for demonstratives (*ese niño/esa niña*) or the plural cases of definites and indefinites. Only for the singular definite and indefinite determiners (*un caso/una casa*) is there additional phonetic information beyond vowel quality, provided the following word starts with a consonant (i.e., *una gata* but not *una abuela*). Table 4 shows the frequencies of all determiner types classified by number in the noun. The frequencies of the more informative singular definite and indefinite forms are indicated in bold.

**Table 4.** Distribution of DPs by determiner and number (all tokens).

Determiner	Plural	Singular	Total
def	68	1011	1079
dem	35	52	87
indef	20	279	299
null	94	535	629
poss	13	292	305
quant	15	27	42
wh	10	7	17
All DPs	255	2203	2458

If we consider only the determiners marked for gender (definite/indefinite/demonstratives) for nouns ending with word markers -a, -o, we are left with 1156 or 47% of the nouns. On the assumption that a speaker or listener would ignore vowel quality for non-alternating nouns ([bok@] for *boca*) and only control production/perception of contrasting vowels, where it matters ([niñ@] would have to be resolved into *niño* or *niña*; similarly for [kas@] *caso/casa*), we could restrict our attention to only alternating nouns. As shown in Table 5, only 791 tokens or 32% of all data has nouns that will mark gender and are accompanied by a determiner with visible gender.

**Table 5.** Frequencies of determiner types associated with gendered nouns (reported for all gendered nouns and separately for only those that entered into an alternation).

Determiner	All Gendered Nouns		Only Alternating Nouns	
	Plural	Singular	Plural	Singular
Definite	50	838	10	651
Indefinite	15	184	6	111
Demonstrative	25	44		13

So far, we have only talked about form. The scenario is even worse if we take into account the specific context of this narrative and deem as highly detectable nouns that are not contextually predictable. As pointed out above, if someone murmurs *Caperucito roj@* we are likely to hear *Caperucita roja* for the simple reason that *Caperucita* is one and only. In our story of choice, there are five main characters; all of them highly contextually predictable. We know we are talking about an *abuela*, not an *abuelo*, and that there is no *loba* in the story. We first note that there is a strong association between semantic type of noun and contextual status, as shown in Table 6. This is in the direction one would expect, with the five main characters (animate, the wolf and humans, *Caperucita*, grandmother, mother and hunter) which make up the bulk of the contextually-given reference. Table 6 cross-tabulates all nouns; columns classify all nouns by whether there was a preceding article or not, and rows show how those nouns are distributed in relation to contextual information, separated by whether it consisted of previous mention, first initial mention of known entities, or neither type of identification. We restricted previous mention to subsequent repetition using the same lexeme; that is, saying *Caperucita* twice was counted as previous mention but not *la niña*).

**Table 6.** Cross-tabulation of semantic type of nouns against contextual status of the DP.

Contextual	Concrete	Abstract	Event	Anim	Hum
Previous mention	19			243	755
Known	2			29	141
Neither	931	163	3	29	129
Grand Total	952	163	3	301	1025

This leaves us with 24 gendered, alternating nouns that were not gender-predictable, neither on the basis of syntax (preceding article) nor context.

### 3.3. Combined Results

The last step in this exploration is to combine what we found about the phonetic realization of these vowels and the morphological analysis. Thus, we will discuss here a way of combining the phonetic results obtained from the narrative elicitation task (the results of the reading task are left aside because read speech is not reflective of the underlying grammatical system) with the analysis of predictability presented in Table 7. Given the characteristics of our dataset (unbalanced number of vowels per category, distribution of predictable and unpredictable nouns both in the story and across speakers), this is very much a tentative proposal rather than a strong claim.

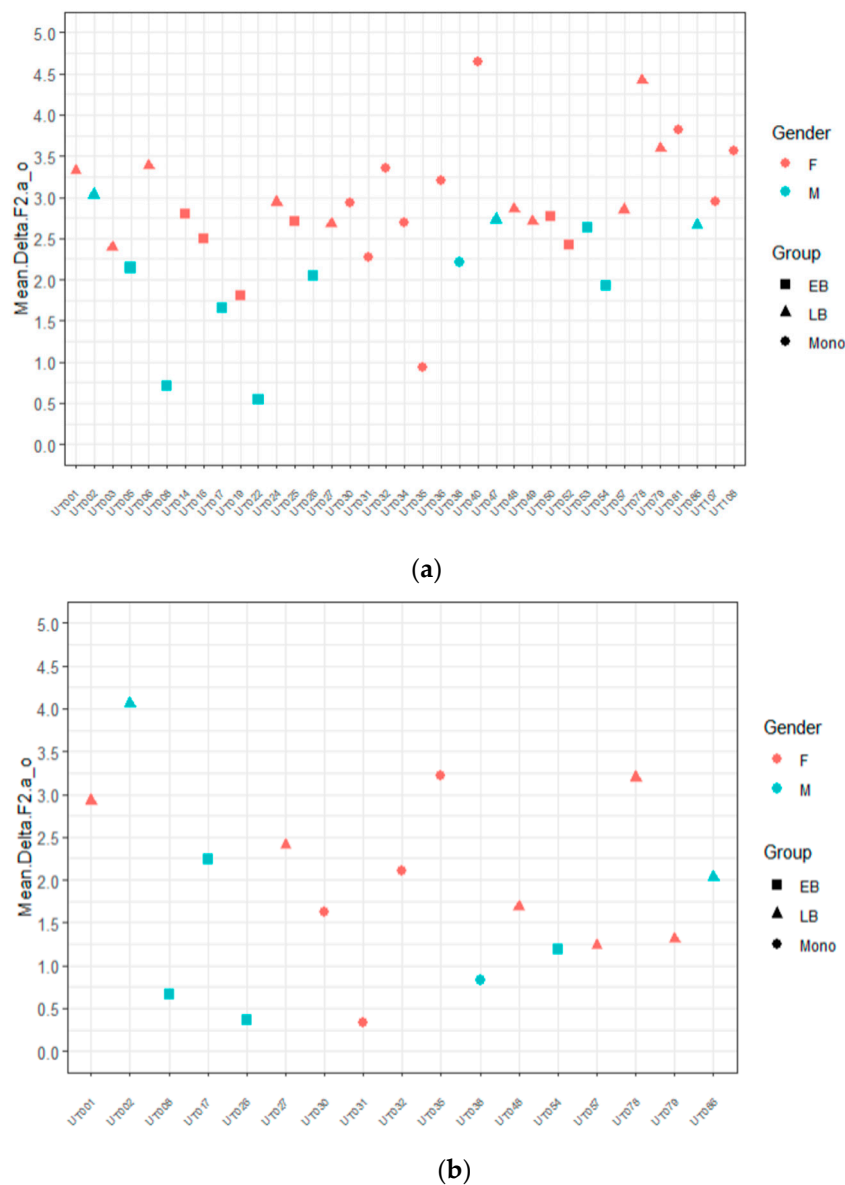
**Table 7.** Cross-tabulation of nouns (reported separately for all nouns and for gendered nouns in alternation) against contextual status of the DP.

Context	All Nouns		Only Gendered Nouns in Alternation	
	Preceding Article	No Syntactic Cues	Preceding Article	No Syntactic Cues
Previous mention	553	465	459	7
Known	80	92	40	1
Neither	669	586	256	24

As discussed in Section 2.3, to combine our results, we Bark-transformed the F1 and F2 values and calculated the F2–F1 difference to obtain a single result for each token. As a reminder, a small difference in Bark is interpreted as a sign of centralization. Our first analysis, displayed in Figures 5 and 6, presents the summary statistics for /a o/.<sup>12</sup> For each vowel, we calculated the mean difference for predictable and non-predictable nouns. Then, to each data point in our Excel file, we subtracted the mean value obtained for the opposite vowel with the same degree of predictability. For example, for a given token of the vowel [a] in *Caperucita*, we subtracted the mean obtained for all the [o] vowels in words like *lobo*.<sup>13</sup> If participants are making no difference between these unstressed vowels, the numbers should be closer to 0.

<sup>12</sup> We focus on these two vowels, given that we had a larger number of tokens than for /e/ and that 80% of the nouns that were marked for gender in the story had these vowels.

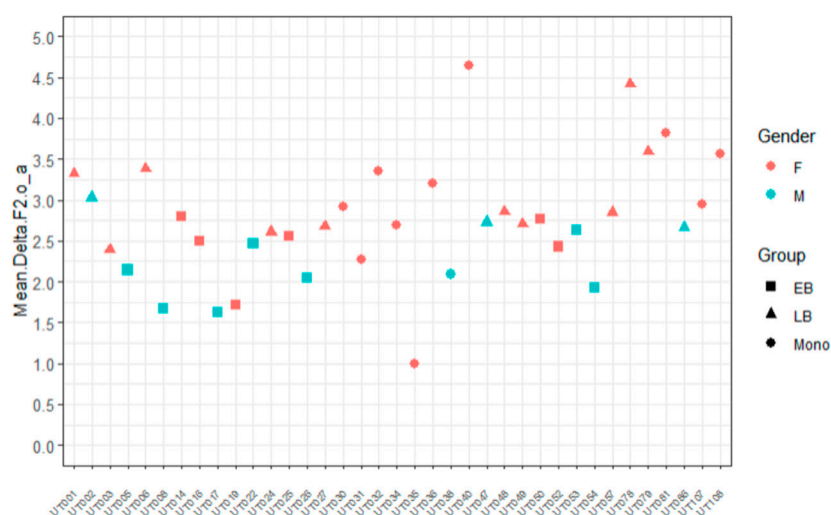
<sup>13</sup> The graphs also reflect the fact that there were many more predictable nouns than non-predictable nouns in this narrative, hence the difference in the number of data points in Figures 5a and 6a when compared to Figures 5b and 6b.



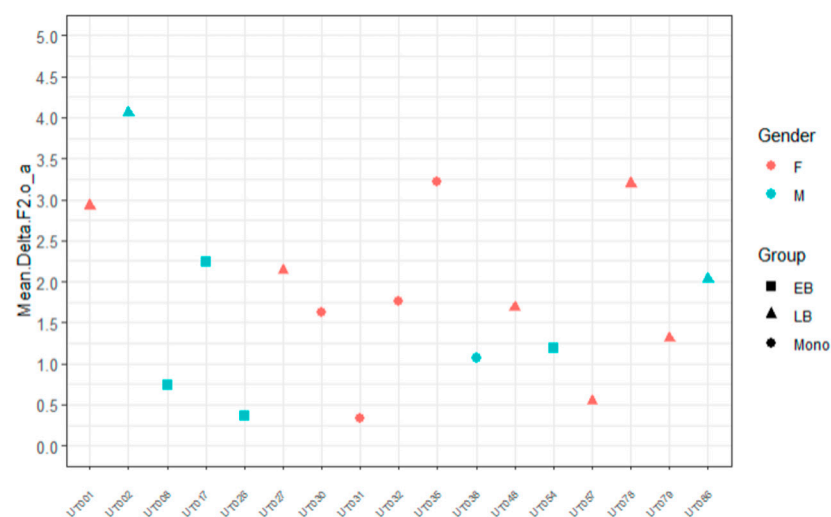
**Figure 5.** F2-F1 in Bark for the vowel /a/: (a) Values obtained for /a/ minus values obtained for /o/ in predictable contexts; (b) values obtained for /a/ minus values obtained for /o/ in non-predictable contexts.

Figure 5 shows that values obtained for EBs are closer to 0 (mean in predictable contexts = 1.95; mean in unpredictable contexts = 0.94) than for the other two groups (LB\_predictable = 3.09; LB\_unpredictable = 2.53; mono\_predictable = 3.02; mono\_unpredictable = 1.99), which means that EBs make less of a distinction between the two vowels. Values obtained for /o/ (Figure 6) display the same tendency, albeit mean values per group are slightly higher than those obtained for /a/.<sup>14</sup> The analysis of both vowels then suggests that the patterns are the opposite to those that we had hypothesized; namely, the difference between the two contrasting vowels was larger in predictable than in non-predictable contexts.

<sup>14</sup> EB\_predictable = 2.15; EB\_non-predictable = 1.33; LB\_predictable = 3.08; LB\_non-predictable = 2.43; mono\_predictable = 2.91; mono\_non-predictable = 1.80.



(a)



(b)

**Figure 6.** F2-F1 in Bark for the vowel /o/: (a) Values obtained for /o/ minus values obtained for /a/ in predictable contexts; (b) values obtained for /o/ minus values obtained for /a/ in non-predictable contexts.

The second type of analysis conducted on this combined dataset was a series of mixed effects models with the F2–F1 (in Bark) difference for each vowel token as the dependent variable and Vowel (/a e o/), Language group (M, EB, LB) and Predictability (yes, no) as the independent variables. Reference levels are /a/ for Vowels, M for Group, and no for predictability. Table 8 reports estimates, standard error, and significance of a linear mixed effects models with Participant as a random effect and Vowel, Language Group, and Predictability of the noun gender as fixed effects (Winter 2020).

The combined analysis of morphosyntactic and phonetic results presented in this section reveals that all groups centralize vowels more in gender-predictable nouns. Further analysis of predictability by formant distance (Figures 5 and 6) suggests that EBs showed a higher degree of centralization in /a o/ than the other groups.



**Table 8.** Results of a linear mixed effects model for the F2–F1 (Bark) with Participant as a random effect and Vowel, Language Group, and Predictability of the noun gender as fixed effects. Reference values: Vowel = [a]; Language Group = Monolinguals; Predictability = no.

Fixed Effects	Estimate	SE	<i>t</i>	<i>p</i> (>  <i>t</i>  )
Intercept	5.05	0.18	27.52	<0.001
Vowel(e)	−3.03	0.12	24.16	<0.001
Vowel(o)	−0.97	0.07	−13.38	<0.001
EBs	−0.10	0.22	0.047	0.63
LBs	−0.11	0.22	−0.50	0.61
Predictability(yes)	0.17	0.09	1.94	0.05

## 4. Discussion

### 4.1. Hypothesis Evaluation

Our first research question and hypothesis targeted the phonetic realization of word-final unstressed singleton vowels and vowel + vowel sequences. We hypothesized that EBs would show a higher rate of overlap in the vocalic spaces of the three target vowels than the other two groups. Results showed this to be the case, particularly in the narrative. Figure 1b and Table 3 point to the fact that the vocalic spaces for the vowel pair that encodes gender more often (i.e., [a]–[o]) overlap in more than 30% of the tokens. This means that 1/3 of these vowels are realized with the same quality. We also hypothesized that monolinguals and LBs would tend to fuse vowels across words while EBs would tend to separate them. Figure 2 showed this to be the case only in the reading task. We further explored the quality (formant trajectories) of the sequences that were pronounced without a pause or a glottal stop to determine potential group differences that could be attributed to influence from English. Given the smaller number of tokens obtained for sequences when compared to single vowels and that most of those tokens were produced by female speakers, we focused on the latter in our statistics for the reading task. We showed that EBs indeed displayed signs of centralization in the realization of such sequences, since /ae/ and /ea/ were realized with similar trajectories and with lower F2 values than those obtained for the other groups. For the narrative, we analyzed both normalized and non-normalized values produced by female and male speakers. Although no statistical differences were observed in the former analysis, the latter revealed signs of centralization in EBs' realization of the sequence /ae/, which was produced with lower F2 values by male speakers and higher F1 values by female speakers than by the monolingual counterparts.

Our second question concerned the perceived errors in gender agreement and concord. Based on previous literature, we expected to find more errors in the EB group than in the other groups. However, very few errors were found in total (only six) and half of them were produced by EBs, which does not seem to support our hypothesis.

Our third research question was explored with data obtained from the narrative and addressed the issue of potential underreporting of gender errors. We hypothesized that vowel centralization would be labelled as accurate gender. Based on the results obtained for the phonetic characterization of vowels (we found 32% of overlap in the vocalic spaces for [a o] in EBs) and the fact that only three errors (examples 7–9) were perceived in this group, we conclude that the hypothesis was confirmed. As we will discuss in the next section, we would have expected at least a higher number of cases labeled as questionable but we ourselves did not do that.

Our final question concerned the results obtained in the combined phonetic and morphosyntactic analysis. Here, we wanted to determine whether EBs would be more likely to mark gender (i.e., both accurate marking and less vowel centralization) in DPs that were not predictable, neither syntactically nor contextually. Our results rejected our hypothesis that this would be the case. Indeed, we found that all groups (and especially EBs) had less centralization in predictable than in non-predictable contexts.

In the next section, we discuss why this may be the case and we contextualize our results in terms of past and future research.

#### 4.2. General Discussion

The results of our phonetic analysis show that there are differences in the realizations across groups and these differences are larger in the narrative than in the reading task. In particular, we have shown that EBs tend to centralize /o/ and have, as a consequence, a high degree of overlap in the vocalic space for [a]–[o]. This is consistent with previous production studies (Gildersleeve-Neumann et al. 2009; Menke 2010; Rogers 2012; Ronquest and Rao 2018) that reported differences in the realization of single vowels between heritage speakers and Spanish monolinguals. The patterns reported here are also consistent with previous perception studies showing that L1 English–L2 Spanish speakers tend to confuse /o/ with /a/ (Morrison 2003; Morrison 2006) and that Spanish heritage speakers tend to confuse vowels but only in unstressed positions (Mazzaro et al. 2016). We have also seen that across-group differences are not restricted to single vowels; there were also differences in the realization of vowels across words. EBs had a higher proportion of hiatuses than the other groups in the reading task and had differences in the F2 trajectories in all the sequences analyzed (i.e., /ae oe ea/). Differences were also found in the narrative for the F1 (female speakers) and F2 (male speakers) trajectories in the /ae/ sequence. Overall, our results showed signs of centralization of the vocalic space, signaled by changes in the F2 rather than in the F1. There are two important points to keep in mind. First, we have found differences in EBs, when compared to the other two groups, in a community with a high level of bilingualism and where Spanish is omni-present. All EBs had or were receiving some education in Spanish and used Spanish daily. They also showed DELE scores very close to those obtained by LBs. Second, we are reporting here results of an unbalanced dataset. Participants are not equally distributed between gender groups (which was overcome in part with normalization) and did not contribute to the sample the same number of tokens (narrative). Although everybody produced the same number of vowels in the reading task, our main analysis refers to the narrative, to which each participant contributed a different number of tokens distributed across different lexical items.

The analysis of gender agreement and concord revealed very few errors, and half of these errors were generated by the early bilinguals. Beyond these few errors, we extracted three observations about the distribution of morphological types. These observations inform our understanding of the challenge of assessing the distribution of gender realizations. First, while most nouns (4 out of 5) are explicitly marked, only a subset of all nouns are accompanied by transparent determiners (3 out of 5). Second, for all transparent nouns and for most transparent determiners, the single cue to gender realization is the vowel: a centralized vowel might actually be ambiguous for accuracy, but vowel centralization is likely to remain undetected, particularly in predictable, non-contrastive contexts. Last, the vast majority of the NPs we analyzed were highly predictable in context and almost never contrastive.

If we now turn to our combined results, we showed in Table 3 that, in 32% of the /a o/ tokens produced by EBs, the vocalic spaces completely overlapped. That means that there should be 221 tokens (out of the 737 tokens of these two vowels) in which the transcriber should have expressed some doubts about the quality of the vowel. If we turn to Section 3.2, we see that we only labeled six cases as gender mismatch. Thus, one can say that, in coding our data, we might have over-reported gender accuracy in possibly as many as 215 tokens, and this only in reference to one of the vowel pairs. If we turn to the combined analysis, both Figures 5 and 6 showed that EBs tend to make smaller distinctions between /a/ and /o/, both in predictable and in non-predictable contexts. To sum up, despite the limitations inherent to this type of data (narrative elicitation) and to our specific dataset (*La Caperucita*), we believe that we have established preliminary grounds for what we have dubbed the *modular interaction hypothesis*, which proposes that changes in one domain (phonetics) can have consequences for the acquisition of other domains (morphosyntax). We have made a case for phonetics as both a methodological obstacle for assessment of bilingual gender grammars as well as a potential incipient factor in contact-induced grammar restructuring in bilinguals.

The results of the combined analysis also showed a main effect of predictability of the noun gender; namely, nouns that were predictable had a significantly higher F2–F1 difference than unpredictable nouns. This means that the distance between the formants, which is a proxy of vowel centralization, was higher in predictable nouns, contrary to our last hypothesis. However, if we return to the data in Table 6, we note that Predictability and Animacy were heavily overlapping categories; i.e., all our predictable nouns were animate, while few inanimate nouns were predictable. This offers an alternative interpretation of the data as signaling that participants are enhancing the marking of gender with animate nouns. Current data does not allow us to evaluate these two possible interpretations of the data.

We do not seek to emphasize the limitations of narrative or spontaneous data, or of previous studies, nor to argue that there are or might be more gender errors than those reported in the literature and/or in the present results. Our goal was to assess how far bilingual speakers can go in reducing the quality of unstressed final vowel contrast before other speakers (including transcribers) detect that something is different or missing in the signal. On the route to this important point, we have shown that, although all DPs are analyzed as possessing abstract gender features, only a portion of them visibly show it. Equally important, only a very small subset of these realized gender forms is not predictable from the syntax or the semantics. As such, these results suggest that future studies of bilingual gender marking should address phonetic analyses explicitly.

The implications are more important than a methodological point. We have examined narrative use of gender and discovered that, in general, gender in many nouns is not overtly marked, but that gender is often a category that can be easily predicted from context. We have also shown evidence, as we predicted, that bilingual heritage adults are likely to produce agreement vowels that are less distinctly specified than those that will be available in the environment of a child in a monolingual community. In a bilingual context with an ongoing contact-based phonetic shift, the signal to gender marking becomes increasingly opaque for the child learner. This contact-induced variation includes robust patterns of vowel centralization and a high degree of overlap in the vocalic space. Furthermore, we should not leave aside the variation that results from the co-articulation between word-final vowels and words beginning with vowels in monolingual speech. Our bilinguals are separating vowel sequences in their production; this means that fused sequences in the speech of monolinguals are likely to remain opaque for them in perception.

Given this variability in the input, we speculate that, at some point, gender in contact varieties of Spanish could become French-like. That is, a scenario where gender is still marked in the syntax but is more visible in the determiner than in the noun morphology. The possibility remains that, for some children, the disruption in the acoustic signal may lead to the development of systems where there has been a reorganization of the underlying featural system, or even a more extreme scenario where gender is no longer part of the agreement system and remains as a vestigial lexical remnant. It is worth recalling the results in [Cuza and Pérez-Tattam \(2016\)](#) data, where, as in other studies, there are more errors with feminine targets than with masculine targets. The individual analysis of their data shows that about a quarter of the children produced no single instance of correct feminine agreement. Those children give the appearance of having contracted the system to the masculine default. Again, more data, particularly data with individual analyses, is needed to explore the range of morphological reorganization processes that might characterize individuals in the North American bilingual context, and to study the potential link between morphosyntax and changes to vowel perception and representation.

Our current study allows us to re-contextualize some of the pre-existing literature. As sociolinguists hold, narratives are a useful tool to elicit semi-spontaneous speech. The narrative analyzed in this paper has special characteristics; namely, it is a traditional folk tale widely known in the Western world. Thus, many lexical items are highly predictable, but not all. In Section 3.2, we showed that there is actually a very small proportion of nouns whose gender is not predictable. This suggests that, in most cases in real life, a bilingual can consistently communicate inserting schwas in word-final position

without them being detected. This, in turn, could be part of the observed asymmetries in results reported in studies such as [Montrul et al. \(2014\)](#), in which participants are accurate in production tasks but inaccurate in grammaticality judgement tasks. The analysis of the narrative, in combination with the study of the phonetic realization of vowels, shows that there are small but systematic differences in the speech of EBs when compared to the other groups. These small differences probably go undetected in production when these speakers interact with other bilingual and monolingual speakers but are substantial enough to limit their perceptual skills and possibly, for some bilingual speakers, prevent them from attaining the underlying representational system.

Our use of semi-spontaneous, narrative data introduced asymmetries and gaps in the data, which precluded certain analyses. At the same time, it gave us a better view of natural speech patterns than the reading data and allowed us to contextualize morphosyntactic analyses and moderate the conclusions from those analyses in light of the phonetic findings.

## 5. Conclusions

We have shown that early bilinguals differ in the phonetic realization of vowels when compared to monolinguals, as has been reported in previous studies. We argued that those differences have implications for the way in which we should analyze agreement errors (or the lack of them) in Spanish-English bilinguals. Even in our population of early bilinguals, who use Spanish daily, we have seen, on the one hand, low rates of reported errors in agreement, as identified by human coders, and on the other, frequent patterns of vowel centralization and a high degree of vowel overlap that cast doubt on informal analyses of gender marking. We argue that such phonetic reality provides the basis for established patterns of incipient restructuring in the morphological system of Spanish heritage bilinguals (as suggested by [Scontras et al. 2018](#)) as well as other, not yet fully documented possible impacts to the system.

Our final take home message is simple. Misperception happens for single segments (e.g., [Calabrese 2012](#); [Durvasula and Kahng 2015a](#); [Durvasula and Kahng 2015b](#); [Hawkins 2010](#); [Repp 1992](#); see also [Ohala 1989, 1993](#)) as well as for whole utterances. Speaking about mondegreens, i.e., misheard song lyrics, [Nevins \(2014\)](#) explains: “listeners might impose what they wish to hear in the songs, or perhaps, what they expect to hear, in perception”. Spanish-speakers predict and expect grammatical gender, so they hear gender vowels when, in fact, the bilingual speaker has spoken a schwa. Thus, it is important to cross-analyze data across linguistic domains and to remember that we linguists, as all human beings, believe we hear what we already know.

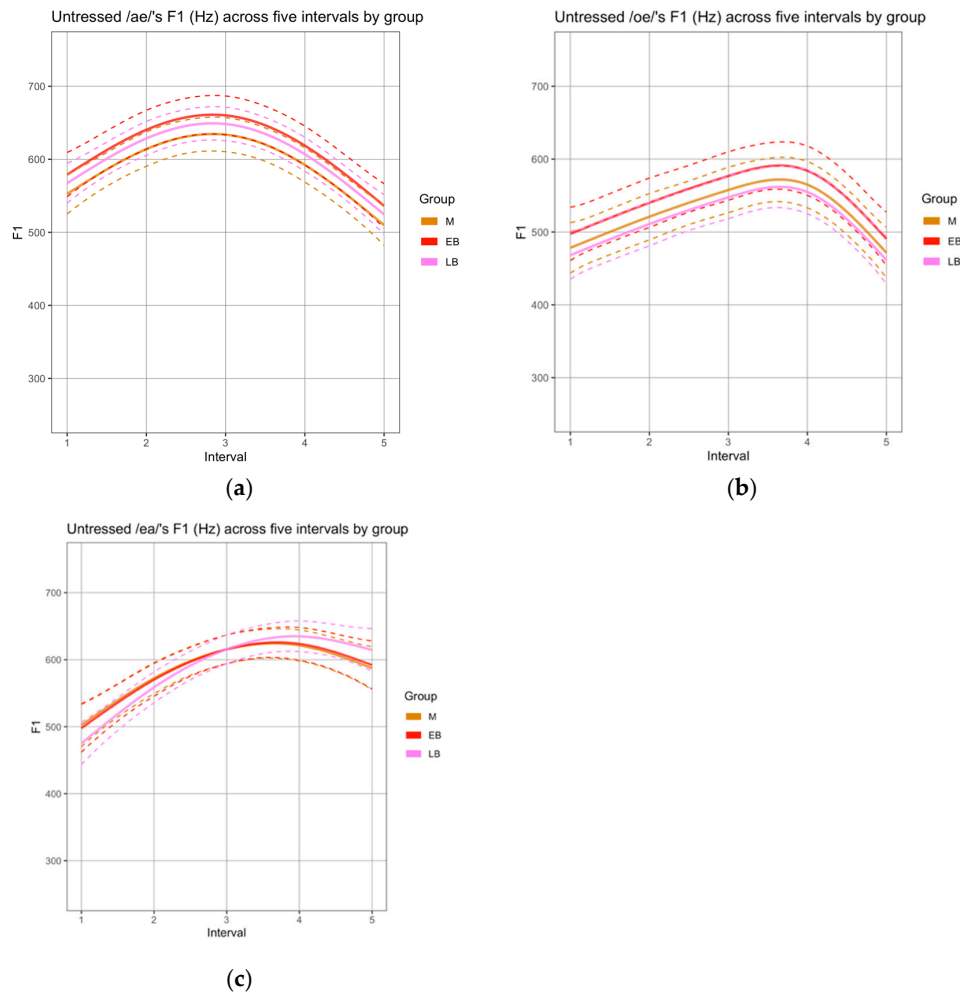
**Author Contributions:** Conceptualization: L.C. and A.T.P.L.; methodology and analyses: all contributors; data collection: N.M.; writing: all contributors; funding acquisition: L.C., N.M. and A.T.P.L. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by a Social Science and Humanities (SSHRC) Insight grant from the Government of Canada, grant number 435-2020-0110.

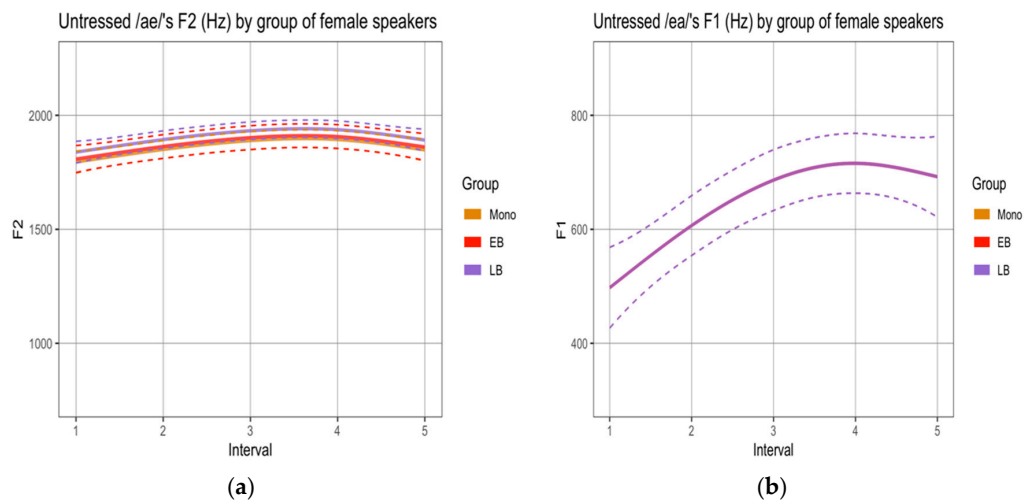
**Acknowledgments:** We thank our participants for their time and M. Barreto and M. Lazzari for their assistance with the statistical analysis.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Appendix A

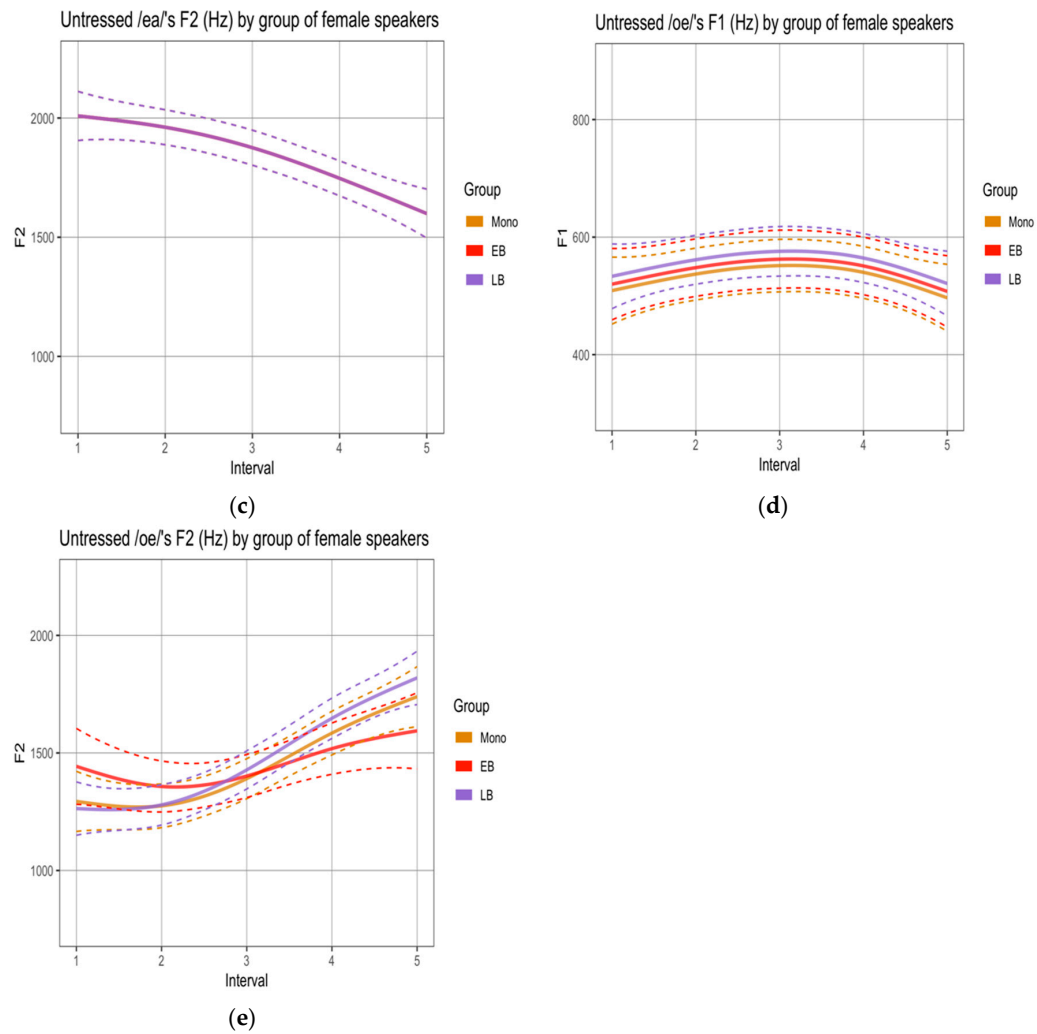


**Figure A1.** SSANOVAs for non-normalized (female) F1 trajectories across five intervals within the sequences (a) /ae/; (b) /oe/; and (c) /ea/ produced by each group of speakers (M, EB, LB) in the reading task.

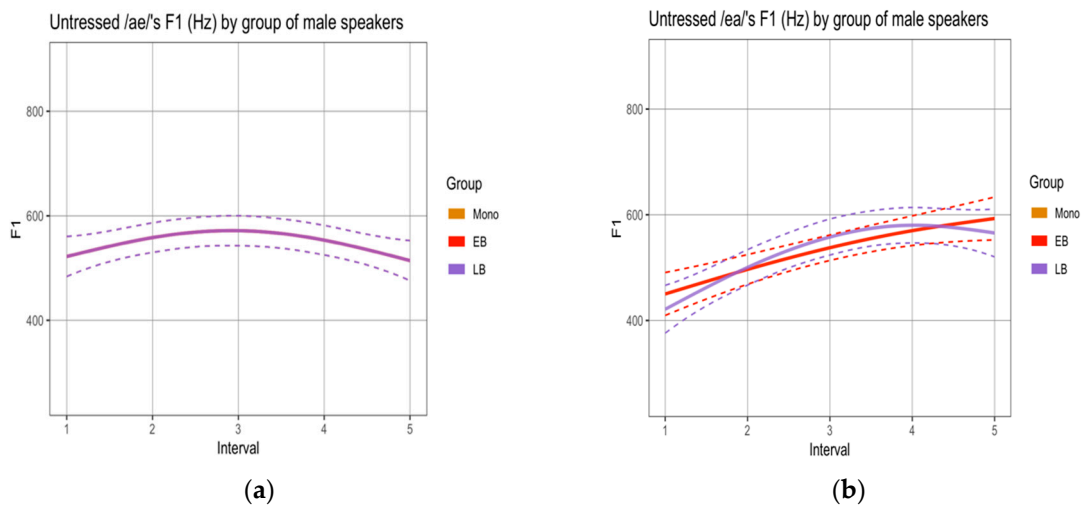


**Figure A2.** Cont.

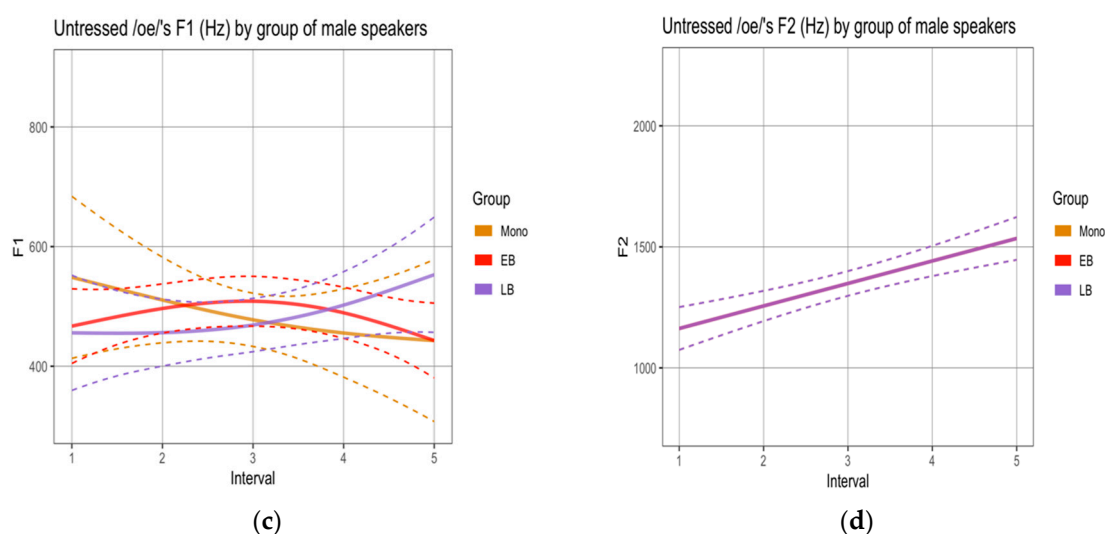




**Figure A2.** SSANOVAs for female fromant trajectories across five intervals within the sequences (a) /ae/ (F2); (b) /ea/ (F1); (c) /ea/ (F2); (d) /oe/ (F1); and (e) /oe/ (F2) produced by each group of speakers (M, EB, LB) in the narrative task.



**Figure A3.** Cont.



**Figure A3.** SSANOVAs for male formant trajectories across five intervals within the sequences (a) /ae/ (F1); (b) /ea/ (F1); (c) /oe/ (F1); and (d) /oe/ (F2) produced by each group of speakers (M, EB, LB) in the narrative task.

## References

- Adank, Patti M., Roel Smits, and Roeland van Hout. 2004. A comparison of vowel normalization procedures for language variation research. *Journal of the Acoustical Society of America* 116: 3099–107. [\[CrossRef\]](#)
- Aguilar, Lourdes. 1999. Differences. *Speech Communication* 28: 57–74. [\[CrossRef\]](#)
- Aguilar, Lourdes. 2010. *Vocales en Grupo*. Madrid: Arco/Libros.
- Alba, Matthew C. 2006. Accounting for variability in the production of Spanish vowel sequences. In *Selected Proceedings of the 9th Hispanic Linguistics Symposium*. Edited by Nuria Sagarra and Almeida. J. Toribio. Somerville: Cascadilla Press, pp. 273–85.
- Barreña, Andoni. 1997. Desarrollo diferenciado de sistemas gramaticales en un niño vasco-español bilingüe. In *Contemporary perspectives on the acquisition of Spanish*. Edited by Ana T. Pérez-Leroux and W. Glass. Somerville: Cascadilla Press, vol. 1, pp. 55–74.
- Bedore, Lisa M., and Laurence B. Leonard. 2001. Grammatical morphology deficits in Spanish-speaking children with specific language impairment. *Journal of Speech, Language, and Hearing Research* 44: 905–24. [\[CrossRef\]](#)
- Bernhardt, Barbara H., and Joseph P. Stemberger. 1998. *Handbook of Phonological Development*. San Diego: Academic Press.
- Boersma, Paul, and David Weenink. 2001. Praat: Doing Phonetics by Computer [Computer Program]. Available online: <http://www.praat.org/> (accessed on 15 June 2020).
- Borzzone de Manrique, Ana M. 1976. Acoustic study of /i, u/ in the Spanish diphthongs. *Language and Speech* 19: 121–28. [\[CrossRef\]](#)
- Calabrese, Andrea. 2012. Auditory representations and phonological illusions: A linguist's perspective on the neuropsychological bases of speech perception. *Journal of Neurolinguistics* 25: 355–81. [\[CrossRef\]](#)
- Calenge, Clément. 2006. The package adehabitat for the R software: A tool for the analysis of space and habitat use by animals. *Ecological Modelling* 197: 516–19. [\[CrossRef\]](#)
- Castilla, Anny P., and Ana T. Pérez-Leroux. 2010. Omissions and substitutions in Spanish object clitics: Developmental optionality as a property of the representational system. *Language Acquisition* 17: 2–25. [\[CrossRef\]](#)
- Colantoni, Laura, and José. I. Hualde. 2016. Constraints on front-mid vowel gliding in Spanish. In *The syllable and Stress*. Edited by R. Nuñez Cedeño. The Hague: Mouton de Gruyter, pp. 1–28.
- Colantoni, L., and A. Limanni. 2010. Where are hiatuses left? A comparative study of vocalic sequences in Argentine Spanish. In *Selected Proceedings of the 38th Linguistic Symposium on Romance Languages*. Edited by En K. Arregi, Z. Fagyal, S. Montrul and A. Tremblay. Amsterdam: John Benjamins, pp. 23–38.

- Culberston, Jennifer, Hanna Jarvinen, Frances Heggarty, and Kenny Smith. 2019. Children's sensitivity to phonological and semantic cues during noun class learning evidence for a phonological bias. *Language* 95: 268–93. [CrossRef]
- Cuza, Alejandro, and Roció Pérez-Tattam. 2016. Grammatical gender selection and phrasal word order in child heritage Spanish: A feature re-assembly approach. *Bilingualism: Language and Cognition* 19: 50–68. [CrossRef]
- Cuza, Alejandro, Ana T. Pérez-Leroux, and Liliana Sánchez. 2013. The role of semantic transfer in clitic-drop among Chinese L1-Spanish L2 bilinguals. *Studies in Second Language Acquisition* 35: 93–125. [CrossRef]
- D'Imperio, Mariapaola, Gorka Elordieta, Sónia Frota, Pilar Prieto, and Marina Viga'rio. 2005. Intonational phrasing in Romance: The role of syntactic and prosodic structure. In *Prosodies*. Edited by Sónia Frota, Marina Viga'rio and Maria J. Freitas. Berlin: Mouton de Gruyter, pp. 59–98.
- Davidson, Lisa, and Daniel Erker. 2014. Hiatus resolution in American English: The case against glide insertion. *Language* 90: 482–514. [CrossRef]
- Delattre, Pierre. 1965. *Comparing the Phonetic Features of English, French, German and Spanish: An Interim Report*. Philadelphia: Julius Groos Verlag.
- Durvasula, Karthik, and Jimin Kahng. 2015a. Illusory vowels in perceptual epenthesis: The role of phonological alternations. *Phonology* 32: 385–416. [CrossRef]
- Durvasula, Karthik, and Jimin Kahng. 2015b. *Phonological Alternations Modulate Illusory Vowels in Perceptual Epenthesis*. Tromsø: Universitetet i Tromsø.
- Eichler, Nadine, Veronika Jansen, and Natascha Müller. 2013. Gender acquisition in bilingual children: French–German, Italian–German, Spanish–German and Italian–French. *International Journal of Bilingualism* 17: 550–72. [CrossRef]
- Ettlinger, Marc, and Jennifer Zapf. 2011. The role of phonology in children's acquisition of the plural. *Language Acquisition* 18: 294–313. [CrossRef] [PubMed]
- Fox, Robert A., James E. Flege, and Murray J. Munro. 1994. The perception of English and Spanish vowels by native English and Spanish listeners: A multidimensional scaling analysis. *Journal of the Acoustic Society of America* 97: 2540–51. [CrossRef] [PubMed]
- Frota, Sonia, Mariapaola D'Imperio, Gorka Elordieta, Pilar Prieto, and Marina Viga'rio. 2007. The phonetics and phonology of intonational phrasing in Romance. In *Prosodic and Segmental Issues in (Romance) Phonology*. Edited by Pilar Prieto, Joan Mascaro and Maria-Josep Sole. Amsterdam: John Benjamins, pp. 131–53.
- Gathercole, Virginia. C. Mueller. 2002. Grammatical gender in bilingual and monolingual children: A Spanish morphosyntactic distinction. In *Language and Literacy in Bilingual Children*. Edited by D. Kimbrough Oller and Rebecca E. Eilers. Clevedon: Multilingual Matters, pp. 207–19.
- Gilbert, Harvey R., Michael P. Robb, and Yang Chen. 1997. Formant frequency development: 15 to 36 months. *Journal of Voice* 11: 260–66. [CrossRef]
- Gildersleeve-Neumann, Christina E., Elizabeth D. Pena, Barbara L. Davis, and Ellen S. Kester. 2009. Effects on L1 during early acquisition of L2: Speech changes in Spanish at first English contact. *Bilingualism: Language and Cognition* 12: 259–72. [CrossRef]
- Goebel-Mahrle, Thomas, and Naomi L. Shin. 2020. A corpus study of child heritage speakers' Spanish gender agreement. *International Journal of Bilingualism* 2020: 1–17. [CrossRef]
- Goldstein, Brian A., and Karen E. Pollock. 2000. Vowel errors in Spanish-speaking children with phonological disorders: A retrospective comparative study. *Clinical Linguistics and Phonetics* 14: 217–34.
- Grinstead, John, Myriam Cantú-Sánchez, and Blanca Flores-Ávalos. 2008. Canonical and epenthetic plural marking in Spanish-speaking children with specific language impairment. *Language Acquisition* 15: 329–49. [CrossRef]
- Gu, Chong. 2014. Smoothing Spline ANOVA Models: R Package Gss. *Journal of Statistical Software* 58: 1–25. Available online: <http://www.jstatsoft.org/v58/i05/> (accessed on 1 July 2020).
- Guasti, Maria. T., Anna Gavarro, Joke De Lange, and Claudia Caprin. 2008. Article omission across child languages. *Language Acquisition* 15: 89–119. [CrossRef]
- Hawkins, Sarah. 2010. Phonological features, auditory objects, and illusions. *Journal of Phonetics* 38: 60–89. [CrossRef]
- Haynes, Erin F., and Michael Taylor. 2014. An assessment of acoustic contrasts between short and long vowels using convex hulls. *Journal of the Acoustical Society of America* 136: 883–91. [CrossRef]
- Hualde, José I. 2014. *Los sonidos del Español*. Cambridge: Cambridge University Press.

- Hualde, José I., Miquel Simonet, and Francisco Torreira. 2008. Postlexical contraction of nonhigh vowels in Spanish. *Lingua* 118: 1906–25. [CrossRef]
- Hutchinson, Sandra P. 1974. Spanish vowel sandhi. In *Parasession on Natural Phonology of the Regional Meeting of the Chicago Linguistic Society*. Chicago: Chicago Linguistic Society, pp. 752–62.
- Johnson, Daniel E. 2015. Quantifying Overlap with Bhattacharyya's Affinity. *Paper presented at NWAV 44, Toronto, ON, Canada, October 25*. Available online: [https://danielezrajohnson.shinyapps.io/nwav\\_44](https://danielezrajohnson.shinyapps.io/nwav_44) (accessed on 28 August 2020).
- Kehoe, Margaret M., and Conxita Lleó. 2017. Vowel reduction in German-Spanish bilinguals. In *Romance-Germanic Bilingual Phonology*. Edited by Mehmet Yavas, Margaret M. Kehoe and Walcir C. Cardoso. London: Equinox, pp. 14–37.
- Kent, Raymond D., and Ann. D. Murray. 1982. Acoustic features of infant vocalic utterances at 3, 6 and 9 months. *Journal of the Acoustic Society of America* 72: 353–65. [CrossRef]
- Kuchenbrandt, Imme. 2005. Gender acquisition in bilingual Spanish. In *Proceedings of the 4th International Symposium on Bilingualism*. Edited by James Cohen, Kara T. McAlister, Kellie Rolstad and Jeff MacSwan. Somerville: Cascadia Press, pp. 1252–63.
- Ladefoged, Peter. 2001. *Vowels and Consonants: An Introduction to the Sounds of Language*. Oxford: Blackwell.
- Larranaga, Pilar, uijarro-Fuentes, and edro. 2012. Clitics in L1 bilingual acquisition. *First Language* 32: 151–75. [CrossRef]
- Lleo', Conxita, and Katherine Demuth. 1999. Prosodic constraints on the emergence of grammatical morphemes: Crosslinguistic evidence from Germanic and romance languages. In *Proceedings of the 23rd annual Boston University Conference on Language Development*. Edited by Annabel Greenhill, Heather Littlefield and Cheryl Tano. Somerville: Cascadia Press, pp. 407–18.
- Lopez Ornat, Susana. 1997. What lies in between a pre-grammatical and a grammatical representation? In *Contemporary Perspectives on the Acquisition of Spanish, Volume 1: Developing Grammars*. Edited by Ana T. Pérez-Leroux and William R. Glass. Somerville: Cascadia Press, pp. 3–20.
- Lobanov, Boris M. 1971. Classification of Russian vowels spoken by different listeners. *Journal of the Acoustical Society of America* 49: 606–8. [CrossRef]
- Mak Brian, Etienne Barnard. 1996. Phone clustering using the Bhattacharyya distance. In *Proceeding of Fourth International Conference on Spoken Language Processing. ICSLP '96*. Philadelphia: IEEE, vol. 4, pp. 2005–8.
- Martínez-Gibson, Elizabeth. A. 2011. A comparative study on gender agreement errors in the spoken Spanish of heritage speakers and second language learners. *Porta Linguarum* 15: 177–93.
- Mateu, Victoria E. 2015. Object Clitic Omission in Child Spanish: Evaluating Representational and Processing Accounts. *Language Acquisition* 22: 240–84. [CrossRef]
- Mazzaro, Natalia, Laura Colantoni, and Alejandro Cuza. 2016. Age effects and the discrimination of consonantal and vocalic contrasts in heritage and native Spanish. In *Romance Linguistics 2013: Selected Proceedings of the 43th Linguistic Symposium on Romance Languages*. Edited by Christina Tortora, Marcel den Dikken, Ignacio L. Montoya and Teresa O'Neill. Amsterdam: John Benjamins, pp. 277–300.
- McCloy, Daniel R. 2016. phonR: Tools for Phoneticians and Phonologists. R Package Version 1.0–7. Available online: <http://drammock.github.io/phonR/> (accessed on 25 June 2020).
- Menke, Mandy. 2010. Examination of the Spanish vowels produced by Spanish-English bilingual children. *Southwest Journal of Linguistics* 28: 98–135.
- Miller, Karen, and Cristina Schmitt. 2010. Effects of variable input in the acquisition of plural in two dialects of Spanish. *Lingua* 120: 1178–93. [CrossRef]
- Mohagheh, Mercedeh. 2016. Connected Speech Processes and Lexical Access in Real-Time Comprehension. Unpublished Ph.D. dissertation, University of Toronto, Toronto, ON, Canada.
- Montrul, Silvina. 2011. Morphological errors in Spanish second language learners and heritage speakers. *Studies in Second Language Acquisition* 33: 155–61. [CrossRef]
- Montrul, Silvina, and Kim Potowski. 2007. Command of gender agreement in school-age Spanish-English bilingual children. *International Journal of Bilingualism* 11: 301–28. [CrossRef]
- Montrul, Silvina, labakova, and oumyana. 2003. Competence similarities between native and near-native speakers: An investigation of the preterite/imperfect contrast in Spanish. *Studies in Second Language Acquisition* 25: 351–98. [CrossRef]

- Montrul, Silvina, Rebecca Foote, and Silvia Perpiñán. 2008. Gender agreement in adult second language learners and Spanish heritage speakers: The effects of age and context of acquisition. *Language Learning* 58: 503–53. [CrossRef]
- Montrul, Silvina, Justin Davidson, Israel de la Fuente, and Rebecca Foote. 2014. Early language experience facilitates the processing of gender agreement in Spanish heritage speakers. *Bilingualism: Language and Cognition* 17: 118–38. [CrossRef]
- Morgan, Gareth P., M. Adelaida Restrepo, and Alejandra Auza. 2013. Comparison of Spanish morphology in monolingual and Spanish-English bilingual children with and without language impairment. *Bilingualism: Language and Cognition* 16: 578–96. [CrossRef]
- Morrison, Geoffrey S. 2003. Perception and production of Spanish vowels by English speakers. In *Proceedings of the 15th International Congress of Phonetic Sciences*. Edited by Maria-Josep Sole', Daniel Recasens and Joaquin Romero. Adelaide: Causal Productions, pp. 1533–36.
- Morrison, Geoffrey S. 2006. L1 and L2 Production and Perception of English and Spanish Vowels: A Statistical Modelling Approach. Unpublished Ph.D. dissertation, University of Alberta, Edmonton, AB, Canada.
- Navarro Toma's, Tomás. 1970. *La pronunciación del español*. Madrid: CSIC.
- Nearey, Terrance M. 1977. Phonetic Feature Systems for Vowels. Ph.D. dissertation, University of Alberta, Edmonton, AB, Canada.
- Nevins, Andrew. 2014. Misheard lyrics Introduction: ABBA's Super Trouper. Available online: <https://youtu.be/dBnhkwRmYuQ> (accessed on 10 October 2020).
- Nicoladis, Elena, and Kristan Marchak. 2011. Le carte blanc or la carte blanche? bilingual children's acquisition of french adjective agreement. *Language Learning* 61: 734–58. [CrossRef]
- Ohala, John J. 1989. Sound change is drawn from a pool of synchronic variation. In *Language Change: Contributions to the Study of Its Causes*. Edited by Leiv Breivik and Ernst Jahr. New York: Mouton de Gruyter, pp. 173–98.
- Ohala, John J. 1993. The phonetics of sound change. In *Historical Linguistics: Problems and Perspectives*. Edited by Charles Jones. London: Longman, pp. 237–78.
- Pérez-Pereira, Miguel. 1991. The acquisition of gender: What Spanish children tell us. *Journal of Child Language* 18: 571–90. [CrossRef] [PubMed]
- R Studio Team. 2020. RStudio: Integrated Development for R. RStudio, PBC, Boston, MA. Available online: <http://www.rstudio.com/> (accessed on 25 May 2020).
- Repp, Bruno H. 1992. Perceptual Restoration of a "Missing" Speech Sound: Auditory Induction or Illusion? *Perception & Psychophysics* 51: 14–32. Available online: <http://myaccess.library.utoronto.ca/login?url=https%3A%2F%2Fsearch.proquest.com%2Fdocview%2F58231001%3Faccountid%3D14771> (accessed on 28 August 2020).
- Rogers, Henry. 2000. *The Sounds of Language: An Introduction to Phonetics*. London: Longman.
- Rogers, Brandon M. 2012. Vowel Quality and Language Contact in Miami Cuban Spanish. Ph.D. dissertation, Brigham Young University, Provo, UT, USA.
- Romanelli, Sofía, Andrea Menegotto, and Ron Smyth. 2018. Stress-induced acoustic variation in L2 and L1 Spanish vowels. *Phonetica* 75: 190–218. [CrossRef] [PubMed]
- Ronquest, Rebecca. 2016. Stylistic variation in heritage Spanish vowel production. *Heritage Language Journal* 13: 275–96. [CrossRef]
- Ronquest, Rebecca, and Rajiv Rao. 2018. Heritage Spanish phonetics and phonology. In *The Routledge Handbook of Spanish as a Heritage Language*. Edited by Kim Potowski and Javier Muñoz-Basols. Abingdon: Routledge, pp. 164–77.
- Rvachew, Susan, Elzbieta. B. Slavinski, Megan Williams, and Carol L. Green. 1996. Formant frequencies of vowels produced by infants with and without early onset otitis media. *Canadian Acoustical Society* 24: 19–28.
- Rvachew, Susan, Karen Mattock, Linda Polka, and Lucie Me'nard. 2006. Developmental and cross-linguistic variation in the infant vowel space: The case of Canadian English and Canadian French. *Journal of the Acoustic Society of America* 120: 2250–59. [CrossRef]
- Schnitzer, Marc L., and Emily Krasinski. 1994. The development of segmental and phonological production in a bilingual child. *Journal of Child Language* 21: 585–622. [CrossRef]
- Scontras, Gregory, Zuzanna Fuchs, and Maria Polinsky. 2015. Heritage language and linguistic theory. *Frontiers in Psychology* 6: 15–45. [CrossRef]



- Scontras, Gregory, Maria Polinsky, and Zuzanna Fuchs. 2018. In support of representational economy: Agreement in heritage Spanish. *Glossa: A Journal of General Linguistics* 3: 1–28. [\[CrossRef\]](#)
- Shoemaker, Ellenor, and Rebekah Rast. 2013. Extracting words from the speech stream at first exposure. *Second Language Research* 29: 165–83. [\[CrossRef\]](#)
- Silva-Corvalán, Carmen. 2014. *Bilingual Language Acquisition: Spanish and English in the First Six Years*. Cambridge: Cambridge University Press.
- Snyder, William, Ann Senghas, and Kelly Inman. 2001. Agreement morphology and the acquisition of noun-drop in Spanish. *Language Acquisition* 9: 157–73. [\[CrossRef\]](#)
- Song, Jae Y., Megha Sundara, and Katherine Demuth. 2009. Phonological constraints on children's production of English third person singular-s. *Journal of Speech, Language, and Hearing Research* 52: 623–64. [\[CrossRef\]](#)
- Stoel-Gammon, Carol, and Karen Pollock. 2009. Vowel development and disorders. In *The handbook of Clinical Linguistics*. Edited by Martin J. Ball, Michael R. Perkins, Nicole Müller and Sara Howard. Hoboken: Blackwell, pp. 525–548.
- Stoel-Gammon, Carol, and Anna Vogel Sosa. 2008. Phonological development. In *Language Development*. Edited by Erika Hoff and Marilyn Shatz. Oxford: Blackwell, pp. 238–56.
- Strelluf, Christopher. 2016. Overlap among back vowels before /l/ in Kansas City. *Language Variation and Change* 28: 379–407. [\[CrossRef\]](#)
- Vokic, Gabriela, and Jorge M. Guitart. 2009. On the realization of non-high Spanish vowels in contact within sequences of more than three vowels: Spectrographic evidence and theoretical implications. *Southwest Journal of Linguistics* 28: 71–102.
- Willis, Erik W. 2005. An initial examination of southwest Spanish vowels. *Southwest Journal of Linguistics* 24: 185–98.
- Winter, Bodo. 2020. *Statistics for Linguists: An Introduction Using R*. New York and London: Routledge.

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).