Listening to Accented Speech in a Second Language: First Language and Age of Acquisition Effects

Saioa Larraza

CNRS. Laboratoire Psychologie de la Perception, UMR 8158
BCBL. Basque Center on Cognition, Brain and Language. Donostia-San Sebastian, Spain

Arthur G. Samuel

BCBL. Basque Center on Cognition, Brain and Language. Donostia-San Sebastian, Spain
IKERBASQUE, Basque Foundation for Science
Department of Psychology, Stony Brook University

Miren Lourdes Oñederra

University of the Basque Country, Vitoria-Gasteiz, Spain
Department of Linguistics and Basque Studies

Corresponding author:

Saioa Larraza

Laboratoire Psychologie de la Perception, UMR 8158

CNRS - Université Paris Descartes

45, Rues des Saints-Pères

75006 Paris

France

Tel. +33(0)1.42.86.43.14

Email: saioa.larraza@parisdescartes.fr

Abstract

Bilingual speakers must acquire the phonemic inventory of two languages and need to recognize spoken words cross-linguistically; a demanding job potentially made even more difficult due to dialectal variation, an intrinsic property of speech. The present work examines how bilinguals perceive second language (L2) accented speech and where accommodation to dialectal variation takes place. Dialectal effects were analyzed at different levels: An AXB discrimination task tapped phonetic-phonological representations, an auditory lexical decision task tested for effects in accessing the lexicon, and an auditory priming task looked for semantic processing effects. Within that central focus, the goal was to see whether perceptual adjustment at a given level is affected by two main linguistic factors: bilinguals' first language and age of acquisition of the L2. Taking advantage of the cross-linguistic situation of the Basque language, bilinguals with different first languages (Spanish or French) and ages of acquisition of Basque (simultaneous, early, or late) were tested. Our use of multiple tasks with multiple types of bilinguals demonstrates that in spite of very similar discrimination capacity, French-Basque versus Spanish-Basque simultaneous bilinguals' performance on lexical access significantly differed. Similarly, results of the early and late groups show that the mapping of phonetic-phonological information onto lexical representations is a more demanding process that accentuates non-native processing difficulties. L1 and AoA effects were more readily overcome in semantic processing; accented variants regularly created priming effects in the different groups of bilinguals.

Key-words

L2 accented speech, highly proficient bilinguals, representation, speech perception, spoken word recognition.

Introduction

Even though spoken words adopt many different acoustic forms based upon the speaker and the circumstances, successful communication is usually achieved. Bilingual contexts and differing dialects further increase sound diversity and phonetic variation. Consequently, the process of dealing with variation can be particularly complicated for the bilingual listener.

Basque bilingualism provides an unusually interesting case because of its geographic and resulting linguistic pattern. Because the Basque language cohabitates with Spanish in one region and with French in another, the phonemic inventories of these two major languages can potentially affect the Basque sound system differently in the two regions. Given that Basque monolingualism only exists in childhood, the Basque region offers a case of "double bilingualism" – Spanish-Basque and French-Basque bilinguals who share one language, and differ in their other language. In the current study we examine how Basque bilingual listeners with different linguistic backgrounds recognize Basque words that include phonetic variation that is attributable to a particular Basque dialect. We provide separate tests to determine whether such variation impacts basic phonetic-phonological processing, lexical access, and/or semantic activation. These tests allow us to identify the locus of any costs incurred by dialectal variation.

The literature offers abundant evidence that two factors may impact second language (L2) processing at one or more representational levels: the listener's native language, and the age of acquisition of L2. Linguistic experience clearly plays a critical role in how well a listener can deal with phonetic variation. Adaptation to such variation has been shown to be better for native listeners than for non-native ones (Garcia Lecumberri, Cooke, & Cutler, 2010; Tuinman, Mitterer, & Cutler, 2012). The native language is one of the principal factors affecting the perception of a second language; the literature is full of examples showing that L1 can hinder the learning of

phonemic contrasts in L2. For instance, Dutch listeners have difficulty distinguishing the English /æ/-/ɛ/ contrast (Weber & Cutler, 2004), the British English /ɪ/-/i/ contrast is hard for native speakers of Spanish (Escudero & Boersma, 2004) and Spanish dominant Catalan proficient bilinguals have trouble perceiving the /ɛ/-/e/ Catalan contrast (Navarra, Sebastian-Galles, & Soto-Faraco, 2005; Sebastian-Galles, Echeverría, & Bosch, 2005). The impact of L1 seems to depend on phonetic salience. For example, the same Spanish-Catalan proficient early bilinguals who had trouble discriminating the /ɛ/-/e/ contrast were able to discriminate the /s/-/z/ Catalan contrast (Sebastian-Galles & Soto-Faraco, 1999).

Another factor that has received substantial attention on second language processing is the age of acquisition (AoA) of L2. Although there are different attitudes towards the value of AoA as a predictor of L2 proficiency, nobody denies that it affects the perception and production of a non-native language. The literature provides varying estimates of this developmental pattern. The classic "critical period" idea was that language learning before around age 12 was more native-like. However, more recent work has generally shown earlier differences. A number of studies suggest that native-like performance requires acquisition by age seven (Caramazza, Yeni-Komshian, Zurif, & Carbone, 1973; Dufour, Kriegel, Alleesaib, & Nguyen, 2014; Kim, Relkin, Lee, & Hirsch, 1997; Silverberg & Samuel, 2004; Sundara & Polka, 2008) or even before age three (e.g., Pallier, Colomé, & Sebastian-Galles, 2001; Samuel & Larraza, 2015; Sebastian-Galles et al., 2005; Sebastian-Galles, Rodríguez-Fornells, de Diego-Balaguer, & Diaz, 2006). Some of the language abilities that have been shown to depend on age of acquisition include phonetic-phonological processing (e.g., Sebastian-Galles & Soto-Faraco, 1999), phonetic-to-lexical mapping (e.g., Samuel & Larraza, 2015), and performance on different syntactic parameters

(Zawiszewski, Gutiérrez, Fernández, & Laka, 2011) or morphosyntactic rules (Johnson & Newport, 1989).

Our focus in the current study is on how well listeners cope with L2 dialectal variation. We test whether the listener's L1 and AoA of L2 determine how L2 dialectal variation is accommodated during phonetic encoding, lexical access, and semantic activation. Prior evidence indicates that this is more successful when an L2 phonological process that occurs in casual speech also operates in the listener's native language (Mitterer & Tuinman, 2012). Our study takes advantage of the fact that the same L2 (Basque) is paired with different L1s (Spanish and French) within a relatively small geographical region, where it is not uncommon to find bilinguals who started learning Basque at different ages.

The Phonemic Contrast of Interest

The phonemic contrast of interest here is a distinction in voiceless alveolar sibilant fricatives that Standard Basque makes contrasting the active place of articulation[1]: apical alveolar /s̻/ vs. laminal alveolar /s̺/. Both sounds of the contrast have the same passive place of articulation (i.e., alveolar), with the tongue placed at the alveolar ridge, just behind the gums. The difference is that the apical alveolar (/s̻/) is pronounced with tip of the tongue against the alveolar ridge, just behind the teeth. In contrast, the laminal alveolar (/s̺/) is made with the blade of the tongue against the alveolar ridge as well. As MacLeod (2012, p. 60) describes it, "perceptually, the difference between the apical and laminal /s/ is that the apical [s̻], due to its more posterior place of articulation, sounds more similar to the voiceless alveopalatal fricative /ʃ/ of English".

[1] Some authors propose instead that the difference between the apical and laminal alveolar fricatives lies in the tongue shape, and there is even discussion about how "grooved" or "sulcal" these shapes are (Catford, 1977; Ladefoged & Maddieson, 1996).

The difference in the active place of articulation between the apical and laminal alveolar Basque sounds is considered to be a one-phonetic-feature contrast that yields minimal pairs in Basque. Critically, this contrast does not exist in Spanish or in French. Nevertheless, the contrast is present in the bilinguals' environment from the beginning of their linguistic experience because in the region where Basque is spoken, bilingualism is always present, exposing listeners to the contrast of interest. However, in one Basque region listeners are less likely to hear this contrast: In the Western dialect – mainly spoken in the province of Biscay – the distinction is neutralized in favor of the apical pronunciation (the side of the contrast present in Spanish, the major language spoken in that region). Neutralization is a frequent type of regular variation with important implications for spoken word recognition. Misperceptions can occur when a contrast is neutralized in one dialect, potentially creating confusions. The Western dialect has additional distinctive morphological and phonological characteristics, but in the current study we specifically focus on only the neutralization.

Table 1 summarizes the distinction that Standard Basque[2] makes that Spanish and French do not. Notice that as in the Western accent, both Spanish and French have only one of these phonemes in their inventories (/s̻/ in Spanish and /s̪/ in French). As Table 1 illustrates, an apicalized articulation of a Basque word containing the laminal fricative could be a word produced with the Western accent of Basque, or it could be a mispronunciation by a Spanish native speaker with L2 Basque. When such speakers learn Basque, they often have difficulties with the pronunciation of the new laminal category not present in their L1, and many such speakers produce it as apical. As the non-native Basque community is quite large, these mispronunciations are widespread. Hence,

[2] There is no unique Standard pronunciation, but the 1998 norms by the Royal Academy of the Basque Language Euskaltzaindia recommend the distinctions found in maximally contrasting dialectal varieties (i.e., Central and Eastern Basque dialects).

when we specify *accented speech,* we are referring to either Western (Basque vernacular pronunciation) or Spanish-influenced pronunciation. In both cases, there is a change from /s̻/ to /s̺/. The specific context we create in each of the experiments will determine whether apicalization comes from Western speakers or Spanish native speakers, and consequently, listeners' treatment of the /s̺/ apical pronunciation.

French-accented Basque does not present any laminalized mispronunciation. As Txillardegi (1980) and Hualde (2010) describe, there is a tendency among French native speakers learning Basque to mispronounce the Basque apical /s̺/ sound more as a prepalatal /ʃ/. The richer variety of fricative sounds present in the French inventory causes a different mispronunciation of the Basque critical contrast than in Spanish. While there is only one fricative sibilant in the inventory of Spanish spoken in the Basque Country (i.e., /s̺/), French includes four such sounds because it distinguishes alveolar and prepalatal[3] places of articulation contrasting in voice (i.e., /s/-/z/ and /ʃ/-/ʒ/).

INSERT TABLE 1 HERE

The goal of the current study is to probe different representational levels to determine at what point variation gets normalized and thus how listeners understand accented forms. Each of the three experiments targets a different possible level: An AXB discrimination task taps phonetic-

[3] Other terms such as *palato-alveolar* or *post-alveolar* fricatives are also used to describe these sounds. We use prepalatal because this is the one that most frequently appears in descriptions about Basque sounds (Hualde, 2010; Oñederra, 2009).

phonological representations, an auditory lexical decision task tests for effects in accessing the lexicon, and an auditory semantic priming task looks for more implicit effects that arise once a lexical representation has been at least partially activated. The lexical decision task is most sensitive to processes involved in mapping the acoustic-phonetic signal to a lexical representation, and the semantic priming task measures the degree to which semantic connections have been made. Collectively, the three experiments provide a systematic investigation of the effects of dialectal variation on bilinguals' perceptual processing, with a focus on how these effects may vary at different representational levels.

Preliminaries

The current study is a relatively large project that has elements of both a standard laboratory-based study and field work. Before presenting the three experiments, it is necessary to provide information about the different populations that were tested, the nature of the critical phonemic contrast being studied, and the ways in which the study mixed field work and laboratory procedures. The three experiments were run within-subjects. All participants first did the auditory semantic priming task, then the AXB discrimination task, and finally, the auditory lexical decision task. This order was chosen to prevent listeners from becoming aware of the phonemic contrast of interest that we presented in the lexical decision task. For clarity of exposition, we present the experiments in ascending order of presumed processing: phonetic, lexical, and semantic association.

Populations

The three experiments in this study were conducted with five groups of participants. It is important to note that even though the participants included in the study were familiar with the neutralization phenomenon of the Western dialect, they did not have this dialect themselves, nor had they been continuously exposed to it. Three of the five groups comprised native Spanish speakers, and two comprised native French speakers. All participants were 18 years of age or older, and spoke Standard Basque very fluently. In a structured interview conducted by professional research assistants who used a standardized scoring system, their average Basque spoken proficiency was 4.3 on a five-point scale. A score of 4 applies to highly fluent speakers of Basque who are able to talk about a wide range of topics, making occasional errors during long and difficult sentences. This is consistent with the participants' self-perceived proficiency of 8.4 on a ten-point scale. All participants were required to complete a detailed linguistic questionnaire that allows us to characterize them very accurately. Among other things, we asked them about their and their parents' native language(s) and native dialect(s); the age of acquisition of each of them; current daily exposure to each of the languages (conversations, TV, radio, classes, etc.); reading/listening/speaking/writing percentage in each language; time spent living abroad, language used for arithmetic, etc. Spanish-Basque bilinguals typically did not speak any French, and the knowledge of Spanish that French-Basque bilinguals had was very basic – most were just able to say a few Spanish words. This pattern is consistent with the fact that even though people lived relatively close to another region, their normal activities were conducted in their own region (and thus in their own language(s)). All reported that they had normal hearing, and all received payment for their participation.

Data collection for most of the participants (83 of the 131) followed standard laboratory-based procedures. The participants were recruited from a pool of subjects at the Basque Center on

Cognition, Brain and Language (BCBL) and tested at the laboratory in sound shielded cabins. The three L1 Spanish groups were (1) Simultaneous bilinguals who learned both Spanish and Basque from birth (N=29; 20 female); (2) Early bilinguals who learned Basque beginning around age three (N=29; 16 female); and (3) Late bilinguals who learned Basque beginning around age seven (N=23; 13 female). All Spanish-Basque bilinguals were originally from Navarre and Gipuzkoa, two of the Basque provinces on the Spanish side of the border. One limitation was the absence of native French speakers who were late learners of Basque – there simply was not a large enough population of such individuals to construct a test group. Thus, we tested two French-Basque groups, Simultaneous bilinguals (N=29; 21 female) and Early bilinguals (N=21; 17 female). These bilinguals all came from Labourd, a Basque province on the French side of the border.

When participants were not able to come to the lab, the recruitment and testing followed procedures that were more of a cross between standard laboratory studies and field work. Some of the participants were recruited via Basque schools and/or universities where they did a part of their studies. This was the case for 6 Spanish-Basque simultaneous, 9 Spanish-Basque early, 4 Spanish-Basque late, 17 French-Basque simultaneous and 12 French-Basque early bilinguals. Participants were tested in quiet rooms at such schools, using the same test programs and headphones used for the subjects tested at the lab, but run on a laptop rather than a desktop PC.

As noted above, two potentially important population factors were varied in assessing accommodation of dialectal variation in a listener's second language. The first such factor was included to test whether the relationship of L1 to L2 affects accommodation at one or more of these processing levels. To assess this, we compared simultaneous bilinguals whose first language was either Spanish or French (henceforth, L1 Spanish and L1 French). Current exposure to Basque was similar for the L1 Spanish and the L1 French simultaneous groups, with both groups reporting

that they used Basque about 50% (sd 16.4) of the time and their other L1 about 50% (sd 17.5) of the time (each language was used approximately equally often in daily conversations, classes, reading, TV/radio, etc.). In both the Spanish and French speaking areas, different varieties of Basque habitate together; vernacular Basque and Spanish- or French-accented Basque. Nevertheless, speakers on each side are not often exposed to the other major language, meaning that Spanish-Basque bilinguals do not listen to French on a daily basis, and French-Basque bilinguals do not hear much Spanish.

The second factor was age of Basque acquisition, to test whether this affects how listeners handle dialectal variation in Basque. For the L1 Spanish subjects, we compared simultaneous to early to late acquisition of Basque. As one would expect, people in the early and late groups tended to use Basque less often in their daily lives than those who learned it from birth. The early group averaged 26% (sd 14.7) usage of Basque, and the late group reported 23% (sd 16.2). For the L1 French groups, we compared the early group to the simultaneous group. The early group reported that they used Basque 36% (sd 12.4) of the time.

Acoustic Properties of the Critical Phonemic Contrast

Because our focus is on how well listeners distinguish the /s̻/-/s̺/ Basque contrast, and do so in light of a dialect that neutralizes the contrast, we conducted acoustic analyses of the Basque apical sound and its Spanish counterpart; different languages can have different instantiations of the same phoneme. We recorded four native speakers of Basque (two female) and four native speakers of Spanish (two female) who did not speak any Basque. The speakers produced the critical sound in intervocalic position in combination with the five vowels that are used in both Basque and Spanish: [as̻a], [es̻e], [is̻i], [os̻o], [us̻u]; each VCV was produced twice. In total, there

were 40 instances of the critical sound in each group. We used Praat software to measure the duration, peak amplitude, and spectral centroid of each token of /s̪/. Figure 1 plots the peak amplitude and spectral centroid for each token. Inspection of the figure suggests that the acoustic realization of /s̪/ in Spanish is more variable than it is in Basque. This may reflect the absence of any similar phonemes in Spanish, allowing it to "spread out" more than it can when it must be distinguished from a similar sound (/s̺/), With respect to specific dimensions, the Basque apical is significantly longer than the Spanish one ($t(39)=3.86$, $p<.001$), is generally a bit louder ($t(39)=2.76$, $p<.01$), and has a lower spectral centroid ($t(39)=3.02$, $p<.005$). This indicates that the acoustic differences between Basque and Spanish /s̪/ sounds could potentially help listeners to interpret this articulation either as a Western neutralization or as a Basque learner's mispronunciation, together with the contextual information where the neutralized sound appears.

INSERT FIGURE 1 HERE

Experiment 1: AXB Discrimination task

As summarized by Weber and Cutler (2004), phoneme discrimination tasks have been used to study non-native listeners' perception of, for instance, English /r/-/l/, /w/-/r/, and /w/-/j/ contrasts by Japanese listeners (Best & Strange, 1992); Zulu click contrasts by English listeners (Best, McRoberts, & Goodell, 2001; Best, McRoberts, & Sithole, 1988); the English /d/-/ð/ contrast by French listeners (Polka, Colantonio, & Sundara, 2001); and German /u/-/y/, /u/-/ʏ/ vocalic contrasts by English listeners (Polka, 1995). In Experiment 1 we measure Spanish-Basque and

French-Basque bilinguals' capacity to distinguish a phonemic contrast that is not present in Spanish or French. We will compare any costs in processing this contrast at the phonetic-phonological level to dialectal variants' effect on lexical access (Experiment 2) and semantic activation (Experiment 3).

We use an AXB discrimination task to learn how bilinguals with different experience with Basque perceive a phonemic contrast that is present in one of their languages, but not in the other. On each trial, a listener hears two acoustically different tokens (A and B) and must decide whether a token heard between them was identical to the first item (AAB) or to the second item (ABB). We test the possible influence that the unshared L1 might have by comparing performance by Spanish-Basque simultaneous bilinguals to performance by French-Basque simultaneous bilinguals. We expect that both simultaneous groups will show very accurate discrimination of the Basque contrast because it is part of one of their native phonemic inventories. To analyze to what extent the AoA of L2 impacts speech perception, we compare bilinguals who differ in the age they started learning Basque. Recall that all participants were highly proficient in Basque. Nonetheless, we expect to find a negative correlation between the age of learning Basque and the ability to discriminate the non-native phonemic contrast. Specifically, early L2 bilinguals should have higher accuracy and/or faster reaction times perceiving the /s̻/-/s̺/ contrast than late L2 bilinguals. In turn, early bilinguals should not perform as well as simultaneous bilinguals. For clarity, we will refer to critical items in terms of their orthography in Basque: The "S" grapheme corresponds to the apical /s̻/ sound and "Z" is used for the laminal /s̺/ sound.

Method

Participants

The five tested groups were described above.

Stimulus Materials

Only nonword stimuli (monosyllabic or disyllabic items) were tested, to avoid any influence of the lexicon and to therefore test purely acoustic-phonetic processing. We embedded the critical sounds in different phonetic contexts that were legal in Basque: #_V (e.g., *za*), V_V (e.g., *azo*), and VC_V (e.g. *ilzo*). We combined the critical sounds with the five Basque vowels (/a/, /e/, /i/, /o/, /u/), in order to cover all sound positions and combinations in which critical phones will appear later in lexical stimuli. Participants listened to 80 critical trials (2 sounds – S and Z – x 20 times each x 2 positions – AAB and ABB); half of the critical trials were monosyllabic and half were disyllabic. The two critical sounds appeared equally often in AAB and ABB trials. 160 control trials were included to provide a comparison to other single-feature contrasts, and to obscure our interest in the critical sounds. All but one of the control phonemic contrasts were legal in Basque, Spanish, and French (e.g., /m/-/n/, /p/-/b/, /t/-/d/, etc.). The /r:/-/r/ control contrast is legal is Basque and Spanish, but not in French. This contrast was presented to all participants but later excluded from all analyses due to its absence in French. Each control trial appeared twice, once as AAB and once as ABB.

Three adult native speakers of Basque (2 male, 1 female) and a fourth adult female highly proficient speaker of Basque produced multiple tokens of the monosyllables and disyllables. All four speakers were also native speakers of Spanish and perfectly distinguished the critical sounds in their productions, as judged by a phonetician who is a native speaker of Basque. They read each token aloud from a randomly ordered list containing four repetitions of each item. The recordings were made in a soundproof booth using a Sennheiser ME65 microphone. The signal was input to

a Marantz PMD-671 digital recorder at a 44.1 kHz sampling rate. Any background noise was then filtered using Goldwave software and individual syllables were extracted using Praat software (Boersma, 2002) by a linguist with phonetic training.

In creating the AXB trials, the "X" item was presented in a male and in a female voice equally often. If it was male, then the "A" and "B" items were presented in two different female voices; if "X" was female, then "A" and "B" were two different male voices. This ensured that the two matching tokens on a trial always were presented in a different voice and gender (e.g., *female1-male1-female2* or *male1-female1-male2*), to force a categorical judgment from listeners, rather than a purely acoustic-physical one. As participants needed to match different exemplars of the same token spoken by different people, they could not make a physical similarity assessment. Rather, they needed to ignore the inter-individual differences in pronunciation and match the middle token to the first or last one in terms of the phonemic information. Participants were explicitly told not to pay attention to any differences in utterances' intonation and to focus on only the segments.

In the Preliminaries section above we reported an acoustic analysis of differences between the apical S in Spanish versus Basque. We conducted a similar acoustic analysis of the apical versus laminal sounds that were tested in the critical stimuli of Experiment 1. Praat software (Boersma, 2002) was used to assess three acoustic properties of the critical sounds. Table 2 shows the resulting measurements for peak amplitude, spectral centroid, and duration. The apical fricative S has a 7 dB higher peak amplitude than the laminal Z, but is about 9 ms shorter. The most robust difference is in the spectral distribution, with laminal sounds generally being higher in frequency than apical ones. Basque apical and laminal fricatives are clearly different in their acoustic characteristics, at least when they are extracted from nonword stimuli and analyzed in isolation.

INSERT TABLE 2 HERE

We also checked the frequency of occurrence of S and Z sounds in the Syllabarium data base for Basque written syllables (Duñabeitia, Cholin, Corral, Perea, & Carreiras, 2010). Across syllables that combine the critical fricatives with the five Basque vowels (i.e., *sa, se, si, so, su* and *za, ze, zi, zo, zu*) the two sounds did not show any significant difference ($t(4)=0.72$, $p=.51$). There are no similar databases for spoken corpora, but given the transparent relationship between orthography and phonology in Basque, these frequencies are a good approximation to those for spoken words.

Design and Procedure

On each trial, listeners heard three tokens. The first (A) and last (B) tokens were always phonetically different, and the middle token (X) was always phonetically identical to either the first one (e.g., *asu-asu-azu*), or the last one (e.g., *za-sa-sa*). The task was to decide whether the middle item matched the first sound (AAB) or the last sound (ABB). Subjects were instructed to press one button with the left index finger when the middle item matched the first sound, and a different button with the right index finger when it matched the last sound. The labels "1" and "3" were shown on the left and right sides, respectively, of a computer monitor.

There were two counterbalanced lists of 240 trials (80 critical plus 160 control – 28 of these control trials were ultimately excluded from all the analyses, given that they were based on the /r:/-/r/ contrast). The lists were identical, except for the voices included in each trial (e.g., *female1-*

male1-female2 or *female2-male1-female1*). Half of the participants were tested with each list. Trials for each listener were individually randomized. All stimuli were presented binaurally over Beyerdynamic DT 770 Pro Studio headphones at a comfortable listening level. The Inter Trial Interval (ITI) was 1000ms, and began after the subject's response. The maximum time participants had to answer before the program presented the next trial was 2500ms. If a response did not occur within this time, the trial was recorded as *no response*. The Inter Stimulus Interval (ISI) was 300ms, between the first and second item, and between the second and third one. Participants were asked to respond as quickly and accurately as possible. To induce the bilinguals to perform in Basque mode, all instructions and all interactions between the experimenter and the participant were in Basque. The task lasted around 12 minutes. The stimulus materials, design and procedure of the experiment were the same for all five subject groups.

Results and Discussion

The central question in Experiment 1 is whether dialectal variants pose any problems for listeners at the phonetic-phonological level. If there is any such cost, we will test whether the cost depends on the two factors in our design: Does having French versus Spanish as an L1 affect perception of the critical contrast, and does its perception depend on when Basque was learned?

Missing responses (1.1%) and trials with responses faster than 200ms or slower than 2500ms were not included in the analyses (0.8%). Due to a software problem, data files were corrupted for some of the subjects, leaving 21 Spanish-Basque simultaneous bilinguals (of 29), 26 Spanish-Basque early bilinguals (of 29), 22 Spanish-Basque late bilinguals (of 23), 24 French-Basque simultaneous bilinguals (of 29), and 17 French-Basque early bilinguals (of 21).

Over the last decade there has been considerable turmoil regarding statistical analysis procedures, with arguments made for linear mixed effects models rather than analysis of variance. Moreover, among those advocating for linear mixed models there is a very strong ongoing debate about best practices. In particular, two very different perspectives exist regarding the benefits and costs of constructing "maximal" models. In the current study, we use linear effects modeling and try to follow the best current practices. On our reading of the literature, this entails building models with factors added until doing so does not improve model fit (Baayen, Davidson, & Bates, 2008; Bates, Kliegl, Vasishth, & Baayen, submitted). As this broad statistical debate plays out in the literature, readers should play close attention to the pattern of the data themselves, as shown in figures and tables.

Analyses were conducted using the lmer function of the lme4 package (Bates, Maechler, Bolker, & Walker, 2013) for linear mixed-effect models in R (R Development Core Team, 2012). For this and for the following experiments, to determine the best-fit model we followed a forward model selection procedure. We started with a null model that only included the random intercepts, and incrementally added predictors and random-effects structure until the fit did not improve. Likelihood ratio tests (Baayen et al., 2008; Dixon, 2008) were used to test whether the inclusion of the fixed and random factors significantly improved the overall fit of the model. The best fitting model for the current data contained both by-subjects and by-items random intercepts and random slopes. L1, AoA of Basque, Condition, and Position were the fixed factors, together with an interaction of L1 by AoA by Condition, given that these provided greater explanatory power to the model. L1 Spanish and L1 French were the levels for L1; control and critical were the levels for Condition; AAB and ABB were the levels for Position; AoA of Basque was introduced as a numerical variable (Appendix A lists all the fixed and random effects included in each model as

well as the lmer formulae used in all analyses). Both accuracy and RT were analyzed as the dependent variables. The reference level for the intercept was set to the L1 Simultaneous Spanish group, control condition and AAB position. For the comparison of different interactions of AoA by L1 by Condition, the reference level was set to the L1 French Simultaneous group. The significance of fixed-effect factors and interactions was evaluated based on the relevant Z-scores and t-tests. For non-Gaussian models, estimations were done based on Z-scores and associated p-values. For Gaussian models, an absolute t-value exceeding 2 was taken as a robust indicator of significance for an alpha level of .05 (Baayen, 2012; Baayen et al., 2008).

Accuracy: Figure 2 shows the accuracy and reaction time results, broken down by Spanish versus French L1, age of acquisition (0, 3, 7), and critical versus control stimuli. Overall, accuracy was higher for control contrasts (97%) than for the critical phonemic contrast (92%) (Intercept: 4.85, SE: 0.21, β: -0.83, SE: 0.26, Wald's z: -3.16, $p<.003$). The interaction of AoA by Condition demonstrates that there was a general tendency among bilinguals to perform poorer on the critical contrast the later Basque was learned (β: -0.12, SE: 0.05, Wald's z: -2.47, $p<.03$). Examination of the means suggests that this effect is primarily being driven by the (L1 Spanish) late bilinguals. Consistent with this view, accuracy was not lower for the French-Basque early group compared to their simultaneous peers; in fact, they were better at discriminating the contrasts than French-Basque simultaneous bilinguals, leading to an interaction of L1 by AoA (β: 0.19, SE: 0.09, Wald's z: 1.99, $p<.05$). However, when the L1 reference level for the intercept was reset to French in order to compare different interactions of L1 by AoA in the language that offers a richer set of L2 ages of acquisition, an age cost arose for the Spanish groups (Intercept: 4.29, SE: 0.22, β: -0.18, SE: 0.09, Wald's z: -1.94, $p=.052$). The lack of interaction of L1 by AoA by Condition shows that the age effect of the Spanish groups was not just for the critical contrast, but also for the control

contrasts (β: -0.14, SE: 0.13, Wald's z: -1.03, p>.05). ABB trials (94%) produced lower performance than AAB trials (96%) did, producing a significant effect of Position (β: -0.57, SE: 0.14, Wald's z: -4.19, p<.001).

Response times: In the analysis of reaction times, incorrect trials (4.73%) were excluded. There was a significant effect of AoA on the reaction times (Intercept: 671.96, SE: 24.5, β: 18.49, SE: 5.49, t: 3.37), with slower responses associated with later AoA of Basque. The age effect was particularly visible on the critical contrast, yielding an interaction of AoA by Condition (β: 8.47, SE: 3.15, t: 2.69). Control contrasts (735ms) were responded to significantly faster than critical ones (790ms) (β: 55.84, SE: 14.89, t: 3.75). In general, French-Basque bilinguals performed significantly slower than Spanish-Basque bilinguals (β: 79.85, SE: 35.36, t: 2.26), regardless of the age they started learning Basque (there was no interaction of L1 by AoA: β: 7.55, SE: 14.45, t: 0.52, *n.s.*). The interaction of L1 by Condition demonstrates that the processing cost on RT for French bilinguals was especially strong for critical contrasts (β: -50.87, SE: 20.78, t: -2.45). ABB trials (738ms) were also significantly faster than AAB trials (772ms) (β: -27.02, SE: 7.43, t: -3.64). The accuracy and RT analyses of Position show a speed-accuracy tradeoff. This would arise if listeners sometimes rapidly responded using a same-different strategy for the first two items in a triplet; ABB pairs would quickly but mistakenly be responded to based on the similarity of the first two members of the triplet. Such rapid responses on AAB trials do not yield errors because the first two members are in fact the same on such trials.

INSERT FIGURE 2 HERE

L1 and AoA effects: The accuracy and reaction time differences for the critical contrasts versus the controls demonstrate that at the phonetic-phonological level of processing, there is a cost for the contrast that is affected by Western dialectal variation. Our more specific tests demonstrate that the two factors tested here – L1 and AoA of Basque – both affect how the critical phonemic contrast is perceived. For the language that allowed the stronger test of AoA, inspection of Figure 2 shows that the Spanish-Basque simultaneous group was the most accurate and fastest, followed by early bilinguals who were about as accurate but a bit slower. Late bilinguals had difficulties distinguishing the S-Z contrast in nonword stimuli; their accuracy was around 85% for the critical sounds, whereas their performance was near ceiling for control sounds. Late bilinguals' poorer performance was also reflected in reaction time. It is worth noting that even though the Spanish-Basque late bilinguals are clearly worse than the other groups, their accuracy of 85% shows that they do have a reasonably good ability to distinguish the apical sound from the laminal one. Presumably they can use one or more of the available acoustic differences (i.e., peak amplitude, spectral centroid and/or duration). This moderate performance level must be considered in the context of the relatively easy AXB discrimination task, where sounds to be matched appear in isolation.

French-Basque bilinguals showed a general reaction time cost in comparison to the faster responses by Spanish bilinguals (Intercept: 671.96, SE: 24.5, β: 79.85, SE: 35.36, t: 2.26). This trend is a bit surprising since the French groups could have done the task using their native phonemic categories (i.e., /s̪/ and /ʃ/, with the latter taken as an exemplar of the Basque apical fricative). As the right panel in Figure 2 illustrates, regardless of bilinguals' L1, the response latency in discriminating the critical contrast was affected by the age bilinguals started learning Basque (β: 8.47, SE: 3.15, t: 2.69). Thus, our results converge with other findings that even early

exposure to a language does not produce the same ability to distinguish a non-native phonemic contrast, compared to native discrimination (Bosch, Costa, & Sebastian-Galles, 2000; Flege & MacKay, 2004; Navarra et al., 2005; Pallier, Bosch, & Sebastian-Galles, 1997; Sebastian-Galles & Soto-Faraco, 1999). Our findings demonstrate that L2 AoA plays an important role in the discrimination of sounds at the phonetic-phonological level, with even very early and very proficient bilinguals showing a cost in discrimination latency.

Consistent with our expectations, the results for Spanish-Basque and French-Basque simultaneous bilinguals suggest that the unshared L1 does not prevent native bilinguals from accurately discriminating the Basque contrast of interest. These results indicate that native learners are quite adept at keeping their two phonemic inventories separate in ways that do not seem to be possible later, even when "later" is only an average age of seven (Spanish-Basque late group), or perhaps even three.

Experiment 2: Auditory Lexical Decision Task

In most current theories of spoken word recognition (e.g., TRACE: McClelland and Elman (1986); Shortlist: Norris (1994), Norris and McQueen (2008); Distributed Cohort Model: Gaskell and Marslen-Wilson (1997)), segmental phonemic representations or features (in the case of the latter model) constitute the connection between physical-acoustic information in the speech signal and representations of words in the lexicon (Pisoni & Luce, 1987). It is often assumed that discrimination difficulties for a contrast at the phonetic level will have an impact on word recognition and activation at the lexical level (Cutler, 2012), though some have argued that the recognition of "speech sounds might require a different skill than recognizing words that contain these sounds" (Diaz, Mitterer, Broersma, & Sebastian-Galles, 2012, p. 680).

The AXB discrimination task in Experiment 1 assessed how well the different bilingual groups could distinguish the Basque (L2) apical-laminal contrast. This provides a baseline for the current experiment in which we study how the neutralization of this Basque contrast by the Western dialect is treated at the lexical level. In the current experiment we use an auditory lexical decision task to look for any effects of dialectal variation on speed and/or accuracy of lexical access. Lexical decision is well suited for this because it provides a measure of listener sensitivity to pronunciation variation (Pitt & Szostak, 2012). The goal is to analyze the phonetic-to-lexical mapping and thus, assess the relationship between discrimination at the phonetic level and efficiency of lexical access.

Previous studies have shown that weakness in phonetic processing may become more apparent as tasks require more lexical processing, rather than purely phonetic-phonological competence. This has been found for both early (Sebastian-Galles & Baus, 2005) and late bilinguals (Diaz et al., 2012). For example, we (Samuel & Larraza, 2015) tested listeners like those in the Spanish early group on an affricate contrast that exists in Basque but not in Spanish, and found deficits in lexical access despite native-like performance on an AXB discrimination test. Sebastian-Galles and her colleagues (e.g., Sebastian-Galles & Baus, 2005) have collected a very rich set of findings on early Spanish-Catalan bilinguals that reveal substantial difficulties at the lexical level despite great fluency in L2.

Thus, in Experiment 2 we analyze whether mispronunciations involving the laminal-apical contrast are accepted in the context of lexical representations. In particular, our aim is to study to what extent this is determined by the linguistic experience of the different Basque bilingual groups. We are especially interested in seeing whether bilinguals accept apicalized pronunciation of Z-

words (i.e., Basque words that contain the laminal /s̻/ sound) as valid Basque utterances based on the fact that the Western dialect includes this neutralization.

In Experiment 1, L1 French bilinguals accurately distinguished the Basque S – Z (/s̻/-/s̻/) contrast in nonwords, but in a lexical context they might become more tuned to the laminalized pronunciation of Basque S-words due to the presence of the laminal Z sound in French words. Conversely, for the L1 Spanish group, the presence of the apical S in the Spanish inventory might make them more likely to accept this variant. Thus, we may find that simultaneous bilinguals have a dominant category when listening to words based on the existing phoneme of the unshared language. Any difficulties at the lexical level are likely to be even more evident for bilinguals who learned their L2 later, given that some difficulties were seen for these listeners even under the simplified listening conditions of Experiment 1.

Method

Participants

The same participants were tested as in Experiment 1.

Stimulus Materials

The stimuli included 48 critical items: 12 S-words, 12 Z-words, 12 S-nonwords and 12 Z-nonwords. Critical nonwords were made by exchanging the apical S and laminal Z sounds. This is a one phonetic feature change that produces two different Basque phonemes, just as in English exchanging /p/ and /b/ would produce mispronounced words like **pank* from *bank*, or **bink* from *pink*. We use the linguistic convention of marking such items with "*", even though such mispronunciations may not be as unacceptable as syntactic or morphological violations.

The words that originally contained the laminal Z sound were made S-like (e.g., [a. s̻e.ri.a] ‘the fox’> *[a.s̻e.ri.a]); these will be called S-nonwords. Basque words with the S sound were made Z-like (e.g., [e.s̻al.di.a] ‘the sentence’ > *[e.s̻al.di.a]), producing Z-nonwords.

The S-nonwords are equivalent to the neutralization that occurs in the Western dialect, whereas the Z-nonwords are not directly linked to any dialect. There might be a few geographical locations where speakers make the S > Z laminalized pronunciation, but this is quite infrequent compared to the extensive areas that neutralize the Basque alveolar fricative contrast into the apical one. The rarity of their production leads us to consider Z-nonwords to be “unlicensed” in terms of Basque dialectal characteristics. We will use the term ‘accented variants’ to refer to the utterances that contain the apical S-neutralization, like **aseria*. However, it should be noted that Western variants sometimes require phonological and morphological changes beyond the Z > S neutralization. For instance, even though we used **aseria* for *azeria* ‘the fox’, Western speakers would actually say *aserixa*, making the Z > S change but also adding an epenthetic consonant intervocalically. Given that our goal is to compare the perceptual accommodation that a single dialectal feature evokes in different listeners, we presented ‘accented variants’ rather than full Western variants. This will allow us to see the effect of the most representative dialectal feature, dissociated from additional Western characteristics.

72 control items were also included: we selected 24 control words that did not contain S or Z. To go with the 48 words (24 critical plus 24 control) there were also 48 control nonwords made with single phonetic feature changes, similar to the /s̻/-/s̺/ exchange. Control nonwords involved phonemic contrasts like the control items in Experiment 1 (e.g., /k/-/g/, /m/-/n/, /p/-/b/). As in Experiment 1, the controls allow us to see whether performance only depends on the presence of a single feature change, or if instead, dialect-based changes pose greater problems. If the error

patterns for S- and Z-nonwords are comparable to the error patterns for control nonwords, then acoustic-phonetic similarity is the simplest explanation. But, if the pattern for S-nonwords does not look like the pattern for control nonwords, the difference may be attributed to an adjustment to accent-based variation.

Items were recorded by an adult male native speaker of Basque who produced a clear distinction between the apical and laminal fricatives. He read each token aloud from a semi-randomly ordered list. Words and nonwords were presented separately, to make the reading easier. For nonwords, he kept the same intonation and articulation as for real words, except for the replaced phoneme. Recordings and acoustic analyses of the items were done with the same procedures and equipment as in Experiment 1.

The critical S-words and Z-words were comparable in terms of word frequency and number of syllables (based on the Basque data base of the University of the Basque Country called 'Euskal Hiztegiaren Maiztasun Egitura' (EHME), http://www.ehu.es/ehg/ehme/). The mean frequency for Z-words was 39.8/million and for S-words was 26.2/million ($t(46)=0.53$, *n.s.*). Both sets of critical words had the same average length, 3.96 syllables.

As in the AXB task, we measured the acoustic properties of the critical stimuli. Fricatives from S-words were compared with fricatives in their matched Z-nonwords (the latter were made by recording an item in which the original apical sound was replaced with the laminal sound). Similarly, fricatives in Z-words were contrasted with fricatives in their matched S-nonwords. The aim of the acoustic analysis was to check whether these items were acoustically distinguishable after the fricative exchange. As Table 3 shows, S-words and Z-nonwords were significantly different in peak amplitude and spectral centroid; the apical sound in S-words is louder and has a

lower centroid. This was also the case for the comparison between Z-words and S-nonwords. Duration was not a useful distinguishing cue in either comparison.

INSERT TABLE 3 HERE

The patterns in Table 3 are very similar to those found in the acoustic analyses of S and Z sounds in Experiment 1's nonword stimuli (Table 2). The difference in the locus of the spectral centroid is a bit smaller here than there, but the effect is still robust. The only change is the now non-significant duration difference, probably due to some over-articulation of nonwords in the stimuli used in the first experiment.

Although lexical stress is not contrastive in Basque (Hualde & Ortiz de Urbina, 2003), we thought it would be prudent to verify that there were no prosodic confounds in the stimuli. Therefore, we analyzed the pitch range variation in critical words and nonwords using Praat. The average minimum pitch for Z-words was 103.2 Hz (*sd* 4.7) and the maximum was 153.5Hz (*sd* 11.3); for S-nonwords the minimum was 103.8 Hz (*sd* 3.6) and the maximum was 148.7Hz (*sd* 8.4). As expected, there was not any significant difference between accented (S-nonwords) and non-accented (Z-words) variants in minimum ($t(23)=0.56$, $p=.58$) or maximum ($t(23)=1.72$, $p=.1$) pitch values.

Design and Procedure

Participants first did a practice of twelve trials made of control items (half words, half nonwords) that were not presented in the real task. These data were not included in the analyses.

On each trial, participants heard a token and the task was to decide whether it was a real Basque word or not. Subjects were instructed to press one button when they considered the token correct and a different button when they thought the token was not a Basque word. The labels 'Nonword' (*Ez-hitza*) and 'Word' (*Hitza*) written in Basque were shown on the left and right sides, respectively, of a computer monitor.

There were two counterbalanced lists of 120 trials (48 critical, 72 control; six control items based on the /r/-/ɾ/ contrast were presented but excluded from the analyses because French does not make this contrast, leaving 66 control items in the analyses) made in such a way that the word and nonword version of the same item never appeared in the same list. Subjects listened to each item once. Half of the participants were tested with each list. Trials for each listener were individually randomized. The Inter Trial Interval (ITI) was 750ms after the subject's response, with a maximum time of 2500ms. *No responses* were treated as in Experiment 1. The task took approximately five minutes.

Participants were informed in advance that nonwords would be very similar to real words, and to only accept a word when they did not notice any mispronunciation in it. In addition, they were told to accept an item as valid if they thought the utterance was something natural that could normally be heard in casual speech conversations. This explicit instruction was designed to make sure that listeners would not be biased against accepting accented variants as valid words. As in any other language, regional accents are expectable when one talks to different Basque speakers, and we wanted this to be true in the experimental context as well. Appendix B lists the stimuli used in the experiment.

Results and Discussion

Both accuracy and RTs were analyzed, with RTs measured from the onset of the item. Model fit was obtained using the same mixed-effect model analysis as in Experiment 1. We conducted statistical analyses using a best fitting model that included by-subject and by-item random intercepts and random slopes. The base model included L1 (Spanish vs. French), Basque AoA (as a numerical variable), Condition (word vs. nonword) and Sound (S, Z and control) as fixed factors. To account for possible confounding interactions among these predictors, the model included a 4-way interaction among the fixed factors: L1 X AoA X Condition X Sound. The reference level was set to the L1 Simultaneous Spanish group, word condition and control sound. Experimental conditions with more than two levels were compared using distinct linear mixed-effect models that only differed in the reference level chosen to estimate the intercept. Missing responses (1.9%) and time-outs (1.2%) were treated as in Experiment 1.

Accuracy: Figure 3 presents the accuracy and reaction time data for the word stimuli, and Figure 4 presents the corresponding data for the nonword stimuli. Overall, accuracy was higher for words (90%) than nonwords (71%) in all groups (Intercept: 4.21, SE: 0.35, β: -1.76, SE: 0.44, Wald's z: -4.02, $p<.001$). As expected, results showed a main effect of AoA (β: -0.12, SE: 0.05, Wald's z: -2.59, $p<.01$); the later Basque was learned, the worse bilinguals performed. L1 French bilinguals performed significantly worse than L1 Spanish bilinguals on the recognition of Basque items containing the Z sound, making more errors both in Z-word acceptance and Z-nonword rejection, as the interaction of L1 by Sound indicates (β: -1.07, SE: 0.42, Wald's z: -2.56, $p<.03$). The S sound created additional difficulty in nonwords, where all groups of bilinguals accepted S-nonwords as valid Basque items, producing a highly significant interaction of Condition by Sound (β: -3.54, SE: 0.58, Wald's z: -6.11, $p<.001$). This is of course due to listeners' perceptual accommodation of the Western accent. In fact, acceptance of this pronunciation is in this sense not

an error. This caveat does not apply to Z-nonwords because there is no dialectal basis for these items. As can be seen in Figure 4, for these stimuli there was a clear negative relationship between the AoA of L2 and the accuracy of Basque lexical access of nonwords with the Z mispronunciation. The significant interaction of L1 x AoA x Condition x Sound confirms that accuracy in Z-nonword rejection decreased the later Spanish-Basque bilinguals learned Basque (Intercept: 2.53, SE: 0.38, β: -0.37, SE: 0.18, Wald's z: -2.05, $p<.05$).

Response times: Incorrect trials (21.6%) were excluded from the reaction time analyses. There was no speed-accuracy tradeoff, as good accuracy was associated with fast responses. A main effect of age of acquisition arose due to the slower RTs for bilinguals with later AoA of Basque (Intercept: 942.05, SE: 26.45, β: 10.30, SE: 5.24, t: 1.97). Even so, the pattern for the latencies in word recognition was extremely similar across all five groups, where Z-words were activated more slowly (1059ms) than control words (981ms) but more quickly than S-words (1111ms). There is a similar pattern in the nonword reaction times. Z-nonwords (1253ms) yielded longer RTs than control nonwords (1156ms), with slowest responding to the critical S-nonwords (1281ms). The consistently slow responses to the S sound yielded a main effect of Sound (β: 114.76, SE: 25.73, t: 4.46). French-Basque bilinguals show an evident age effect in nonword rejection, producing an interaction of L1 by AoA by Condition (β: 40.57, SE: 12.97, t: 3.13). There was also a significant interaction of AoA by Condition by Sound (β: 13.16, SE: 5.32, t: 2.47), indicating that as the age of learning Basque increases, the slower bilinguals are in rejecting Z-nonwords that have no dialectal basis. In contrast, the time needed to recognize S-nonwords did not generally depend on bilinguals' L2 AoA (in this case, there is no interaction of AoA by Condition by Sound: β: -4.75, SE: 6.77, t: -0.71, n.s.). The lack of a 4-way interaction of L1 by AoA by Condition by Sound indicates that the slower responses showed by French bilinguals for

Z-nonword rejection do not depend on L2 AoA (β: 17.59, SE: 13.76, t: 1.28, n.s.). The poorer accuracy and slower reaction times bilinguals showed on both critical sounds relative to the controls confirm that these sounds impair lexical access.

INSERT FIGURE 3 HERE

INSERT FIGURE 4 HERE

Response times to accented variants: To have a better understanding of the time course of the processing of dialectal versus unlicensed variation, here we provide a description of the reaction times for nonwords that were accepted as words (17.4% of the total items). Given the small number of remaining data points, statistical analyses were not applied here; we simply describe the apparent pattern. Figure 5 shows the reaction time results for the two comparisons of interest. The first one, performance on Z-words versus S-nonwords, measures how well each group does in accommodating L2 dialectal variation. The second one, performance on S-nonwords versus Z-nonwords, shows the difference between accepting accented versus unlicensed variants. For the case of most interest, accented variants (e.g., *aseria) took longer to be accessed than their standard counterparts (e.g., azeria 'the fox'). The RT difference between them was in general bigger for Spanish-Basque bilinguals (on average, a bit over 100ms longer) than for French-Basque ones (about 50ms longer). Basque bilinguals with quite different linguistic experiences

accommodate accents while recognizing spoken words, but this has a response time cost. As noted previously, accepted S-nonwords may not be true errors because these items could arise due to the Western dialect, or a Spanish accent.

The three stimulus types produced very consistent responses across the different groups: Accented variants were activated more slowly than Z-words, but faster than Z-nonwords. Z-nonwords that are not licensed by a Basque dialect seem to be the most difficult ones to deal with. When listeners incorrectly accepted them, this decision usually required more time. In the General Discussion, we will consider implications of these results for the time course of perceptual adaptation mechanisms.

INSERT FIGURE 5 HERE

Correlation analyses: Because the same subjects did the AXB task and the lexical decision task, we can test whether basic discrimination ability can account for the rates of inappropriate acceptance of unlicensed variation (i.e., Z-nonwords). We correlated the extent to which bilinguals' performance on Basque nonwords was determined by their discrimination capacity at the phonetic-phonological level. As Figure 6 shows, there is a clear negative correlation between the AXB accuracy results for the critical contrast and the acceptance of Z-nonwords in the lexical discrimination task; the better that listeners discriminate the critical contrast, the less they accept Z-nonwords. The correlations were statistically robust for the three L1 Spanish groups: Spanish-Basque simultaneous (r=-0.69, $t(20)$=4.12, p<.001), Spanish-Basque early (r=-0.66, $t(24)$=4.22, p<.001) and Spanish-Basque late bilinguals (r=-0.56, $t(21)$=3, p<.01). The effect was marginally

significant for the French-Basque simultaneous group (r=-0.4, t(23)=2.02, p=.055), and non-significant for the French-Basque early bilinguals (r=-0.1, t(16)=0.38, p=.71). If we consider the Z-nonwords to be illicit, then listeners should only accept them if they don't notice the change. Bilinguals who have trouble discriminating the two sounds are likely to misperceive the Basque variants, confusing the /s̪/ and /ʂ/ sounds. These listeners should be most likely to not notice the change, and this is exactly what the negative correlations demonstrate.

INSERT FIGURE 6 HERE

Interestingly, the acceptance of S-nonwords was in general not related to basic phonemic discrimination of the critical contrast. Four out of the five tested groups did not show any even remotely significant correlation between the two measures: Spanish-Basque simultaneous (r=0.25, p=.27), Spanish-Basque early (r=-0.11, p=.6), Spanish-Basque late (r=-0.11, p=.62) and French-Basque early bilinguals (r=0.35, p=.17). Only the French-Basque simultaneous group showed a significant correlation (r=0.43, p<.05), and this correlation was different than what one might expect, if a correlation was expected: the better that listeners distinguished the two critical sounds in the AXB task, the more they accepted S-nonwords. It might be the case that those French bilinguals with better discrimination abilities were also the ones with more awareness of dialectal variation. This contrasting pattern of correlations – robust for rejecting unlicensed variation, weak for accommodating dialectal variation – suggests that once a reasonable level of phonetic encoding is reached, the ability to accommodate dialectal variation relies on a different type of knowledge.

This conclusion is supported by the fact that, regardless of L1 or AoA, listeners accepted potentially accented variants about 75% of the time (L1 Spanish 78%, L1 French 75%, Spanish-Basque early bilinguals 66%, French-Basque early bilinguals 82% and the Spanish-Basque late group 71%), a pattern that is totally unlike what we observed for the unlicensed Z-nonwords. Different groups of bilinguals are aware of both the Western neutralization and Spanish-accented pronunciations, allowing these listeners to interpret accented variants as what they are: feasible utterances produced either by Western speakers or Spanish-accented non-native Basque speakers.

L1 and AoA effects: The conditions of Experiment 2 actually could have induced listeners to increase their 'Nonword' responses, for two reasons. First, because the speaker who produced the stimuli was not a Western speaker himself, the S-nonwords were acoustically slightly different from S-nonwords that would have been produced by a native Western speaker (see Appendix C for more details). Second, hearing the same speaker produce both a clear /s̻/-/ʂ̻/ distinction and the /s̻/>/ʂ̻/ neutralization could be interpreted as dialectally incongruent. Despite these two factors, all groups systematically accepted S-nonwords. This indicates that there is a strong tendency among Basque listeners to perceive S-nonwords as real utterances. This result converges with research showing that listeners usually adapt to accents, and furthermore, do so very quickly (Clarke & Garrett, 2004; Floccia, Goslin, Girard, & Konopczynski, 2006). Previous research, with Spanish-Catalan early bilinguals (Sebastian-Galles & Baus, 2005) and with Dutch-English late bilinguals (Diaz et al., 2012), found relatively high competence in tasks with low lexical involvement (e.g., a phonemic categorization task) but lower performance as tasks required more lexical processing. The results of the current study follow this pattern, especially for the Spanish-Basque early group. Even though the absolute accuracy of Spanish-Basque early and simultaneous bilinguals was very similar in the AXB discrimination task, on the word recognition task L1 and AoA effects became

larger, as the overall accuracy and RT differences indicate. Collectively, the data support the view that the age of acquisition of a second language has a stronger impact when listeners have to process the L2 at the lexical level, rather than simply encoding it phonetically.

Comparing the results on the AXB test to those on the lexical decision task also suggests that the effect of the listeners' first language at the phonetic-phonological level is magnified for processing at the lexical level. The French-Basque simultaneous and early groups demonstrated that they were quite able to distinguish the S and Z phones: simultaneous bilinguals performed with 94% accuracy on the AXB task, and the early group achieved 96% correct. However, this capacity was greatly reduced when they operated at the word level, where French-Basque simultaneous (36%) and early bilinguals (41%) accepted Z-nonwords more often than Spanish-Basque simultaneous (22%) or early (25%) bilinguals. It is likely that the unshared L1 is responsible for this. French-Basque bilinguals are listening to French words that contain the laminal Z (but not the apical S sound) on a daily basis. Thus, the native laminal phoneme of French words will still sound familiar enough to create spurious activation of Z-nonwords. This would make the listeners more likely to accept the laminalization of Basque S-words, in a way that someone who is not exposed to French (i.e., Spanish-Basque simultaneous bilinguals) would not. This tendency for an increased acceptance of Z-nonwords is likely have its origin in French –their native language, and not in French-accented Basque (which tends to mispronounce the apical S sound as the palatal /ʃ/, a sound present in French but not in Spanish).

Experiment 3: Auditory Semantic Priming

The first two experiments have clarified how different Basque bilinguals discriminate an L2 contrast at the phonetic-phonological level and how they recognize L2 accented variants during

lexical access. The current experiment continues the lexical level analysis, but now goes beyond initial access to examine semantic activation caused by L2 accented variants. The goal is to determine whether bilingual accommodation of dialectal variation affects lexical processing beyond the mapping from sound to lexical entry. We use a semantic priming methodology to examine this issue.

Semantic priming experiments offer evidence for the associative strength between concepts in semantic memory (Meyer & Schvaneveldt, 1971), showing faster processing when a word (the *target*) is semantically related to a concept previously presented (the *prime*). The effect depends on the interaction between related items in the lexicon. The current experiment is designed to test whether Basque accented variants (i.e., S-nonwords) are as effective as unaccented tokens (i.e., Z-words) in priming a semantically related word. Studies using comparable semantic priming tasks have found that regardless of the native dialect, there is a benefit for variants that are not regionally-marked over dialectal ones (Sumner & Samuel, 2009). Several studies have found weaker lexical activation for reduced pronunciation variants than for their citation forms (LoCasto & Connine, 2002; Pitt, 2009; Ranbom & Connine, 2007; Sumner & Samuel, 2005).

As in the first two experiments, performance will be assessed as a function of the subject's L1 and age of Basque acquisition. Given that both Spanish-Basque and French-Basque simultaneous bilinguals evidenced native-like phonetic encoding and lexical access, we would expect that semantic priming should be comparable for standard and accented variants. Similarly, both French-Basque and Spanish-Basque early bilinguals showed a strong tendency to accept accented words in the lexical decision task and were able to discriminate the Basque fricative contrast. Thus, like the simultaneous groups, these listeners are likely to show normal semantic priming even in the face of dialectal variation. Unlike these groups, the Spanish-Basque late L2

bilinguals showed difficulties perceiving the non-native contrast, and detecting S-Z exchanges at the word level. These listeners may show reduced priming, or they could show comparable priming effects for standard and accented variants for a quite different reason than the other groups – a failure to discriminate the variation could lead to priming by a variant.

In order to manipulate the likelihood of listeners' treating an S-nonword as an accent to be accommodated (as opposed to a non-native mispronunciation), we introduced a Context manipulation in Experiment 3. Priming effects for accented variants will potentially be larger when listeners are placed in a context in which a speaker shows evidence of speaking with a Western accent. The Context factor was included as a between-subjects manipulation, with two levels: Standard and Dialectal. Participants listened to recorded instructions before doing the experiment, and depending on the Context to which they were assigned, these instructions were spoken either in Standard Basque or Western (Dialectal) Basque. In the Dialectal context listeners have more contextual clues to interpret the Z > S articulation as the neutralization of the Western accent rather than as a mispronunciation.

The current experiment provides a test of whether any issues of perceptual accommodation are limited to initial access – the phonetic and lexical processes that were tested in the first two experiments. If adjustment to variation is strictly a matter of initial encoding, then even groups who had problems perceptually adjusting to legal versus unlicensed dialectal variation in the previous tests should show no differential costs in semantic activation.

Method

Participants

The same bilinguals who took part in Experiments 1 and 2 did this task.

Stimulus Materials

A semantic priming experiment requires pairs of words (prime-target pairs) that are semantically related. Because we did not have access to a set of word association norms for Basque, the items had to be validated indirectly. We chose 48 associated word pairs from a Spanish database (http://inico.usal.es/usuarios/gimc/normas/buscar_nal.asp, Grupo de Investigación de Memoria y Cognición (GiMC), the University of Salamanca). We then confirmed that the translations of the selected pairs were also highly associated in English (using the Edinburgh Associative Thesaurus, EAT, http://www.eat.rl.ac.uk/). If a word pair is semantically related in both Spanish (e.g., *mesa-silla*) and in English (e.g., *table-chair*), the corresponding pair of words in Basque (e.g., *mahaia-aulkia*) should be related as well. Thus, the chosen Spanish word pairs were translated into Basque.

Each of the 48 associated word pairs was used to create a stimulus set consisting of four prime-target pairs, representing the four conditions tested in the current experiment: Related, Unrelated, Accented, and Control:

Related primes were Z-words in Standard Basque form (e.g., PRIME: *beta**z**ala* 'the eyelid'-TARGET: *begia* 'the eye') that were semantically related to the target words.

Unrelated primes were matched in frequency (using Basque word-frequency norms in the E-Hitz database, Perea et al., 2006) to the Related prime (e.g., PRIME: *kiroldegia* 'the sports centre' – TARGET: *begia* 'the eye)', where *betazala* and *kiroldegia* match in frequency. In this condition, there was no semantic relationship between the prime and the target.

Accented primes provided a within-item baseline for any priming effects. The prime was the same word used as the prime in the Related condition, but with the Z > S neutralization (as in the Western accent); e.g., **beta**s**ala – begia* 'the eye'.

Control primes included a single-phonetic-feature-mispronunciation of the prime from the Related case, similar to the Z > S single feature change. However, exchanges like /g/ > /k/, /d/ > /t/, /n/ > /m/, /e/ > /i/ were used, rather than the critical laminal-apical change. An example of a control nonword is **bed**azala* (/t/ > /d/) – *begia* 'the eye'. Bigram frequency was taken into consideration when creating the control primes, so that the bigrams that appeared in these items were as frequent as the ones in real words.

The target was kept identical across the four conditions within a stimulus set – only primes varied. The four priming conditions were a within-subjects factor, with four counterbalanced lists created using a Latin square design such that the four primes were rotated among the lists. Each target word was presented to a participant only once. All subjects received the same 48 critical targets. The design thus provides a within-subject and within-item measure of priming. If there is a bigger priming effect for *Related* primes than for *Unrelated* ones, but the priming effect is reduced in the *Accented condition*, it cannot be due to weak relatedness (given the Related-Unrelated difference). The comparison between Accented variants and Control mispronunciations will tell us about possible differences between processing nonwords that have a Basque dialectal cause and those that do not. If an Accented prime is more effective than a Control prime, it would mean that listeners are able to relate an accented variant to the standard variant, indicating perceptual adjustment to the dialectal variation.

As noted above, the Western dialect includes changes beyond the merger that we have focused on. In Experiment 3 we included a manipulation to measure if there are any processing effects that these further dialectal characteristics cause beyond those of the single dialectal feature. To test this, we presented "accented" primes (that only differed from standard variants in the fricative merger) versus full Western primes (that contained the additional dialectal morpho-

phonological features). Thus, we presented two types of Accented items: Half of the trials included a *single-feature-change* – the only difference between the Related and Accented prime was the Western Z > S neutralization. The other 24 trials included 2-*feature-changes*. These were cases in which the Western variants not only include the Z > S change, but also require an additional phonological or morphological characteristic of the dialect in order to be considered real dialectal utterances. Including these items strengthened the dialectal manipulation, measuring any difference in the perceptual adjustment driven by the fricative merger versus full Western variants. Within the 2-feature-change trials, control nonwords contained the same number of changed features as Western variants did, compared to the standard form. Thus, if the Western variant includes the Z > S change plus a vowel dissimilation (e.g., *emastia* for the standard form *emaztea:* 1) vowel dissimilation: *-e* + *-a* > *-ia*, 2) /s̪/ > /ʂ̪/ neutralization), then the control nonword would contain the same changes. Specifically, it would include the same vowel dissimilation as the Western variant, together with a control feature changed relative to the standard variant (e.g., *emaztea* > *enaztia:* 1) *-e* + *-a* > *-ia*, 2) /m/ > /n/).

Table 4 provides an example of the *single-feature-change trials* and an example of the *2-feature-change trials* (see Appendix D for a representative sample of stimuli):

INSERT TABLE 4 HERE

Stimuli were recorded by an adult native female speaker of the Western dialect. This speaker was especially appropriate because she was bidialectal due to having lived in different

regions during her life: She was able to make a clear distinction of the /s̪/-/s̺/ contrast when speaking in Standard Basque, which usually is not the case among Western speakers. She read each token aloud from a semi-randomly ordered list. Standard and Western variants were listed separately in order to facilitate the dialect switching by the speaker. The recordings were made with the same equipment as in Experiments 1 and 2. Individual audio files were created for each word, with no silence left at the beginning or end of each file.

The Z-words that were the basis for items in the Related, Control mispronunciation and Accented conditions were all 3.3 syllables long on average. The critical laminal Z sound (or the neutralized apical sound in the case of the Accented condition) appeared 42% of the time in the first syllable, 48% of the time in the second syllable, 6% of the time in the third syllable, and 4% of the time in the fourth syllable. The contexts for Z were the following: 34% word initial (#_); 24% before a consonant (_C); and 42% intervocalic (V_V). The Unrelated words were 3.6 syllables long on average. None of them had an obvious semantic relationship to the words presented in the other three conditions.

In addition to the 48 critical trials, there were 240 filler trials. These pairs included the same proportion of nonwords and words in prime and target positions as on critical trials. Fillers were included to balance the nonwords included on the critical trials, and to greatly reduce the proportion of related items that each listener heard. The large number of fillers also obscured our interest in the critical sounds. Nonwords were phonotactically legal but meaningless Basque strings. Among fillers, 120 of the primes were words and 120 were nonwords; similarly, 120 of the targets were words and 120 were nonwords. Each nonword used as a filler was made with a single-phonetic-feature change relative to its counterpart standard word. Primes and targets were randomly combined to create filler trials.

Fillers, apart from providing distracters for the primary measurement, were also used as part of the Context manipulation. This Context factor was included as a between-subjects manipulation, with two levels: Standard and Dialectal. Filler words were pronounced in the Western accent for the *Dialectal Context* whereas for the *Standard Context* they were articulated in the Standard accent. Thus, different contexts were distinguished by the pronunciation of the filler words (nonwords were the same in both contexts). As noted above, to reinforce the implicit information given by the Context (either Dialectal or Standard), before doing the task participants listened to instructions recorded in either the Standard or Western dialect. Instructions were recorded by the same bidialectal speaker who recorded the stimuli. We carefully avoided the presentation of the critical Z and S sounds in both versions of the instructions. Thus, when they started the task, participants had not heard any exemplars of the fricatives in question.

Design and Procedure

The task was an auditory lexical decision test with semantic priming. On each trial, participants heard two tokens and they had to decide whether the second one was a real Basque word or not. A prime-target ISI of 500ms was used because similar experiments using this ISI have produced robust priming effects (e.g. Sumner & Samuel, 2005; Zhang & Samuel, 2015). Subjects were instructed to press one response button when the second item (the *target*) was a valid Basque word and a different button when the second item was not a real word. The labels 'Nonword' and 'Word' written in Basque were shown on the left and right sides, respectively, of a computer monitor.

There were four counterbalanced lists of 288 trials (48 critical, 240 filler). The four counterbalanced lists were crossed with the two Contexts, resulting in eight stimulus conditions (4

lists x 2 contexts). Participants were randomly assigned to these eight conditions in such a way that there were approximately the same number of subjects per list and context. Trials for each listener were individually randomized. All stimuli were presented binaurally using the same equipment as in Experiments 1 and 2. Each trial began 1000ms after the subject's response on the previous trial. The maximum time participants had to respond before the program presented the next trial was 2500ms. The task took approximately 23 minutes.

Results and Discussion

Both accuracy and RTs were analyzed. Model fit was obtained using the same mixed-effect model analysis as in Experiments 1 and 2. Statistical analyses were conducted using a best-fit model with both by-subject and by-item random intercepts and random slopes. L1 (Spanish vs. French), Basque AoA (as a numerical variable), Context (Standard vs. Dialect), and Priming Condition (with six levels: Related, Unrelated, Accented_1-feature change, Accented_2-feature change, Control Nonword_1-feature change, Control Nonword_2-feature change) were the fixed factors, and the interaction of L1 by AoA by Condition was also included. The reference level was set to L1 Simultaneous Spanish, Standard Context and the Unrelated condition. We compared experimental conditions by using distinct linear mixed-effect models that differed in the reference level chosen to estimate the intercept.

Participants with accuracy lower than 85% on the lexical decision task were not included in the analyses, eliminating two subjects from the Spanish early group, one subject from the late Spanish group, and four from the French-Basque simultaneous group. One subject was excluded from the Spanish-Basque simultaneous group for failing to understand the task. The remaining sample sizes were 28 subjects for the Spanish-Basque simultaneous group, 26 subjects for the

Spanish early group, 22 subjects for the Spanish late group, 26 for the simultaneous French group, and 20 French early bilinguals. Missing responses (0.6%) and trials with responses faster than 200ms or slower than 2500ms (0.7%) were excluded from the analyses.

Table 5 shows the accuracy and reaction time results for the remaining data, and Figure 7 shows the corresponding priming effects.

INSERT TABLE 5 HERE

Accuracy: Overall accuracy was above 93% correct for all of the priming conditions. Performance on 1-feature Accented and 1-feature Control Nonword primes was essentially at ceiling, with slightly worse accuracy for the primes with 2-feature changes. Lexical decision was in most cases less accurate in the Unrelated condition than in the Related case, or than in the conditions with primes derived from the Related case. The accuracy differences were larger between the Unrelated and 1-feature Accented primes, which produced a marginal effect of Condition (Intercept: 4.73, SE: 0.58, β: 4.24, SE: 2.2, Wald's z: 1.93, p=.054; the z values for the rest of the conditions were between 0.9 and 1.7, with the p values between .37 and .08). The two L1 French groups responded with very similar accuracy, while Spanish-Basque bilinguals performed worse the later they started learning Basque. The absence of any age effect for L1 French bilinguals, and the presence of one for the L1 Spanish groups, resulted in a marginal interaction of L1 by AoA (Intercept: 4.38, SE: 0.61, β: 0.39, SE: 0.21, Wald's z: -1.83, p=.06). Overall, performance on each of the priming conditions was not affected by L1 and AoA of L2 (all the interactions of L1 by AoA by Condition were non-significant, with z values between -0.1

and 0.4, and p values between .6 and .87). There was no Context effect (β: 0.33, SE: 0.24, Wald's z: 1.38, p>.05), indicating that the Dialectal context did not influence recognition accuracy of the targets.

Response times: Incorrect trials (3.1%) were not included in the RT analyses. As expected, all groups were on average faster responding in the Related condition (924ms) than in the Unrelated one (960ms), leading to a marginal effect of Condition (Intercept: 914.69, SE: 28.65, β: -36.41, SE: 24.01, t: -1.7). Primes that minimally differed from the Related condition also produced responses that were significantly faster than those in the Unrelated case: 1-feature Accented (895ms, β: -74.6, SE: 28.37, t: -2.63) and 1-feature Control Nonword (905ms, β: -56.91, SE: 28.51, t: -2.1). The primes that included a 2-feature change did not differ from the Unrelated case (*t* values between -0.1 and -0.8). In addition, as Table 5 clearly indicates, the later a bilingual started to learn Basque, the slower they were to respond to the target words, resulting in a significant effect of AoA (β: 13.23, SE: 5.59, t: 2.37). And given that the age of starting to learn Basque equally affected bilinguals with different native languages, there was no interaction of L1 by AoA (β: -2.74, SE: 14.75, t: -0.19, n.s.). Similarly, there was a lack of interaction of L1 by AoA by Condition, as all the priming conditions showed a comparable L1 and AoA effect (the t values for all conditions were between -1 and 0.98).

Given evidence for rapid adjustment to non-standard pronunciations in priming tasks (Witteman, Weber, & McQueen, 2013), we examined whether the patterns of priming changed over time. An analysis comparing the priming effects created by accented primes in the first half of the experiment was not different from the corresponding priming in the second half. Considering the accuracy and reaction time data together, it is clear that performance is dominated by phonetic similarity of primes to the Related case: The Related primes, and ones that only differ by one

feature, enhance performance, whereas primes that are more distant phonetically behave more like the Unrelated case.

As with accuracy, Context did not affect the RT results (β: -10.71, SE: 22, t: -0.49, *n.s.*). Collectively, the results show that explicit exposure to the Western dialect in the Dialectal context did not facilitate semantic association by variants presented in the task.

INSERT FIGURE 7 HERE

Figure 7 summarizes the reaction time priming effects in each of the conditions, relative to the Unrelated case. The figure highlights the three central results of our test of dialectal variation's impact on semantic activation. First, the 1-feature-change Accented condition generally produced results quite like those for the 1-feature-change Control Nonword condition, and both of these were similar to the Related condition. As in the current study, prior work in the literature shows that 1-feature-change primes usually produce significant priming (Broersma & Cutler, 2008; Connine, Blasko, & Titone, 1993; Marslen-Wilson & Zwitserlood, 1989). Connine and her colleagues have argued that spoken words are mapped onto stored lexical representations based on goodness-of-fit. This hypothesis and the current results suggest that lexical representations of spoken words allow a minimal mismatching between the input and stored knowledge, particularly in higher level tasks that do not require focused attention on low-level details.

The second central result is that the first three conditions (Related, and single-feature changes) produced larger priming effects than those in the 2-feature change cases, indicating that priming primarily depends on phonetic similarity. This pattern suggests that unlike the process of

lexical access, the influence of a dialect in activating semantically associated entries is much more subtle. The relatively small priming effect for the 1-feature-change conditions for the late group (see Figure 7) is consistent with the results of Silverberg and Samuel (2004), who found that the age of acquisition of L2 has a strong effect on semantic priming effects, with weaker effects in late bilinguals.

Correlation analyses: Recall that when we examined the correlation between listeners' basic phonetic processing of the critical contrast and their lexical decision performance, we found a sharp dissociation between dialectal variants and unlicensed variants: Discrimination ability was clearly linked to the latter, but unrelated to the former – accommodation of dialectal variation was not grounded in basic phonemic discrimination. We conducted comparable correlation analyses between bilinguals' performance on the AXB discrimination task and the priming effects caused by the 1-feature Accented prime in the current experiment. As we saw with lexical access, basic discrimination ability does not affect semantic activation. None of the groups produced a significant correlation between the two measures: Spanish-Basque simultaneous (r=-0.42, p=.06), Spanish-Basque early (r=0.004, p=.98), Spanish-Basque late (r=0.05, p=.84), French-Basque simultaneous (r=-0.13, p=.59), and French-Basque early (r=-0.25, p=.35). This further suggests that the priming effect by the S-nonwords (i.e. 1-feature-change Accented condition) does not depend on the capacity listeners have to distinguish the two critical sounds, but is instead related to a processing mechanism that enables a match between the accented and standard variant.

L1 and AoA effects: The results of Experiment 3 show single-feature variation does not seem to influence the way one word relates to another. This pattern contrasts with what we observed in the first two experiments, where we saw that L1 and AoA of L2 matter for dealing with variation at a low level, and even more so for initial lexical access. Thus, it makes sense to

distinguish semantic activation during spoken word recognition from activation processes involved in the mapping of sound to lexical entries.

Using long-term repetition priming, Sumner and Samuel (2005, 2009) also found different results for earlier versus later processing (though on a different time scale than the test here). They reported that different dialectal variants are equivalent in the short-term, but in the long-term there is an advantage for regular and canonical forms. The auditory semantic priming task here has shown that functionally speaking, for most of the tested groups, there are generally not significant differences between standard and minimally different accented variants. Nevertheless, this does not mean that these variants are representationally equivalent (Sumner & Samuel, 2009). Language use and linguistic representations may differ in listeners with different linguistic backgrounds, as the detailed comparison of the results in Experiments 1 and 2 suggests. Experiment 3 shows that the fricative neutralization by the Western accent does not alter the way that associated words relate to each other. Because 1-feature-change control nonwords acted similarly, the results may be best thought of within the goodness-of-fit proposal by Connine et al. (1993), where minimally different pronunciations (regardless of being dialectally legal or not) do not affect the results in indirect language processing, as in this auditory semantic priming task. Our results suggest that Basque bilinguals have enough perceptual flexibility to deal with minimally diverging variation. Minimal variation, both dialectally-driven and arbitrary, is well tolerated after initial phonetic recognition and lexical access have been accomplished.

General Discussion

Our use of multiple tasks with multiple types of bilinguals demonstrates that it is important to consider results at different representational levels separately, and to draw specific conclusions

for particular combinations of listeners and tasks when analyzing speech processing. Manipulations that affect phonetic encoding may not affect lexical processing, and vice versa; effects that are relatively small at one level may be magnified at another.

Processing at the phonetic, lexical and post-access levels

The current study, like previous work, demonstrates that bilinguals' first language affects speech perception and spoken word recognition in their second language. The current experiments, more specifically, offer evidence that the native language affects how dialectal variation is treated at each representational level; importantly, the effects created by the native language varied across representational levels. The French sound system is richer than Spanish in the sense that it distinguishes alveolar and prepalatal places of articulation contrasting in voice (i.e., /s/-/z/ and /ʃ/-/ʒ/). The phonemic inventory of Spanish spoken in the Basque Country differentiates neither place nor voice in fricative sibilants, resulting in only one phoneme, /s̪/, against the four French phonemes. We found that this robust difference in fricative inventories did not produce substantial differences in the discrimination performance on the critical Basque contrast by simultaneous bilinguals. In contrast, when given a lexical decision task, French versus Spanish simultaneous bilinguals showed significant differences in lexical access. L1 French bilinguals were both less accurate and slower than their L1 Spanish counterparts. This dissociation between the phonetic-phonological and lexical level shows that the ability to accurately identify the critical segments does not guarantee a high level of performance on spoken word recognition.

It might be the case that listeners confronted by a more complete system – i.e., French – need more time to acquire and establish it than listeners whose native system is simpler. If so, for the early bilinguals, the native system of French listeners would have been less consolidated at the

moment they started learning Basque than the system of native Spanish speakers. Reduced competition from L1 might have facilitated the acquisition of the non-native contrast for French-Basque early bilinguals. Alternatively, native listeners of French might have superior implicit phonetic meta-knowledge due to the richness of their native fricative inventory, enhancing their ability to learn a new phonemic contrast. For the native language Spanish speakers, the age of L2 acquisition had a very large effect, with significant costs in both accuracy and reaction time on the discrimination task. Consistent with this, there was a numerical advantage for the French-Basque early bilinguals compared to the Spanish-Basque early bilinguals in discriminating the critical apical and laminal sounds. AoA effects were more evident for the Spanish speakers, in part because it was possible to include a group of late learners, who were clearly affected by their late acquisition despite being quite fluent in Basque. Our results show clear effects of both the native language and the age of starting to learn Basque, and some ways in which these factors interact.

A clear conclusion from the current study is that the mapping of phonetic-phonological information onto lexical representations is a more demanding process that accentuates non-native processing difficulties (Diaz et al., 2012; Sebastian-Galles & Baus, 2005). There were consistent costs for the critical sounds compared to control sounds when we measured lexical access with the lexical decision task in Experiment 2. These effects were strongly modulated by the age of acquisition of Basque. The AoA results may be best understood in the context of the view that it is not the age of starting to learn L2 *per se* that is critical. Several proposals (Flege, 1999; Iverson et al., 2003; Kuhl, 1994; Sebastian-Galles et al., 2005) instead claim that the neurological stage of L1 development – how established it is at the starting point of learning subsequent languages – is what affects the development of second language listening. It is clear that the longer the experience

with the L1, the more consolidated its internal structure will be. An implicit role for age is assumed, but it is not presumed to be the main causal factor.

Unlike the process of mapping the acoustic-phonetic input to a lexical entry, semantic processing was not affected significantly by dialectal variation. Instead, semantic access was primarily influenced by the phonetic similarity of primes and targets: A single-feature deviation was tolerated whether it was dialectally justified or not, whereas two-feature changes produced reliably weaker priming effects than those caused by truly Related primes or ones that were off by a single feature. Thus, dialectal variation affected lexical access most strongly, basic discriminability moderately, and semantic processing minimally.

Relationships between different processing levels

On the lexical decision task, all groups systematically accepted S-nonwords, the items made by replacing laminal Z sounds with apical S sounds, as occurs in the Western dialect. The high acceptance rate is a clear reflection of perceptual accommodation for accented speech. The neutralization in Basque accented speech drives listeners to readjust their phonetic-phonological representations and thus recognize accented words. Hence, flexible perceivers adjust their lexical representations by matching the apical phonetic surface of accented variants to the underlying laminal phoneme of stored items. This accommodation of dialectal variation did not depend on basic discrimination ability for the critical sounds: The correlation analyses relating performance on the AXB discrimination task to lexical access showed that S-nonword acceptance was not correlated with bilinguals' discrimination capacity. In principle, the accommodation of dialectal variation could be accomplished by storing two representations of each word, one with the standard pronunciation and one with the dialectal neutralization. In a recent study of Spanish-Basque

bilinguals, we (Samuel and Larraza (2015)) found that under conditions in which dual representations were not plausible, listeners mapped two phonetic variants onto a single lexical representation. Given this outcome with similar bilinguals, it seems unlikely that accommodation of the Western dialect is accomplished by storing additional lexical representations; listeners instead appear to adjust to the dialectal input.

In contrast, there is no such adjustment for phonetic variation that is not dialect-based, such as that in the items made by replacing the apical S sound by the laminal Z sound. The acceptance of such Z-nonwords was clearly related to bilinguals' phonemic discrimination ability, as shown in the correlation analyses. A poor ability to distinguish the laminal /s̻/ and the apical /s̪/ sounds by late bilinguals leads to the activation of Z-nonwords, regardless of the (non)existence of these variants in any Basque dialect.

The tasks we chose for Experiments 2 and 3 were designed to tap different aspects of word recognition. The lexical decision task is generally taken as a sensitive measure of the process of making contact with the lexicon, whereas semantic priming tasks reflect processes that follow once the lexicon has been (at least, partially) accessed. Alternatively, we could conceptualize the difference by saying that lexical decision requires explicit recognition of variants, while the priming task is more implicit, in the sense that it relies on indirect processing of activated variants. Across the different groups we tested, when listeners were *explicitly* asked about utterances, they consistently activated accented items more slowly than they did standard words. In contrast, in *implicit processing*, accented variants did not produce any significant cost compared to standard words. Our interpretation is that depending on the adopted attentional set, an on-line normalization mechanism may or may not be applied. Pitt and Szostak (2012) showed that listeners making lexical decisions changed their responses based on whether they were instructed to focus on details

of pronunciation versus lexical content. Listeners were more tolerant of variation in the latter case, which the authors interpreted as an attentionally driven effect.

Explicit decisions about items require focusing on the pronunciation, enhancing any possible mismatch between the input and stored lexical representations. Under these conditions, it is likely that mechanisms to adjust basic phonetic-phonological representations will be applied. When our listeners needed to focus on the pronunciation and make explicit decisions about it, a short-term adjustment process may have operated to solve the phonological mismatch present in the input. Phonological processing determined that the apical pronunciation present in the Western accent was not a context-dependent feature, as it is not attributable to assimilation by the context. Thus, the apicalization can be treated as a contrast merger. An adjustment process of sibilant apicalization is then incorporated into the sound structure of the language in such a way that it enables a phonetic-to-lexical mapping process. This is how bilinguals were able to access and recognize accented variants in a way that was not applicable to other types of mispronunciations, even if that recognition process had a time cost.

When no such explicit decisions were required (in the auditory semantic priming task), listeners did not adopt such a careful approach towards pronunciation details. As a result, variation could be more easily overlooked. From a goodness-of-fit perspective (Connine et al., 1993), variants that were close enough to their counterpart standard words did not incur added costs for integrating the variation. In indirect processing, minimal mismatches between the signal and the stored knowledge are overcome and do not seem to require any normalization. The independence of such indirect processing and basic phonetic analysis is consistent with the correlation analyses that we conducted: Unlike what we saw for lexical access, semantic priming was uncorrelated with listeners' ability to distinguish the Basque critical sounds.

Theoretical implications

One of the central goals of this study was to determine the specific locus or loci where perceptual accommodation to a dialect takes place. There is a growing body of evidence that indicates that perceptual adjustment takes place at the prelexical phonological representational level. Models such as Shortlist (Norris, 1994; Norris and McQueen, 2008) and Merge (Norris, McQueen, & Cutler, 2000) hold that normalization processes operate at such an intermediate prelexical level. The adjustment to variability is achieved through an abstract code that gets linked to the stored word form (McQueen & Cutler, 2001). In these models, mismatches between phonetic forms and the forms stored in the lexicon will initially generate poor lexical activation. However, the mismatch provides a learning signal at the prelexical level, changing the representations at this level so that the representations adapt to the deviant input (Norris, McQueen, & Cutler, 2003; Samuel & Kraljic, 2009). This process cleans up any within-speaker difference (within the same accent) before the phonological representation is used to access the lexicon. Analyzing the mechanisms that underlie perceptual adjustment, Sumner (2011) also suggested that low level changes act in perception via pre-lexical mapping mechanisms. Therefore, the adaptation that accented speech produces in the existing prelexical representations can subsequently improve dialectal processing. These psycholinguistic findings converge with the linguistic claim by (Donegan and Stampe (2009, pp. 31-32)) that the output of phonological processing is not entirely surface or allophonic: "inputs are indeed unlimited, but linguistic perception is phonemic, and as a result […] linguistic 'intentions' are limited". This position asserts that, despite variation in the articulated surface form, listeners will perceive the speech based on their own phonemes, enabling a match between different phonetic variants.

Listeners in the current study were adults living in a region in which they would have had occasional exposure, over the course of their lifetimes, to the Western dialect. Our experiments show that speech perception in such bilinguals is fine-tuned to phonetic details as well as to more abstract phonological and lexical aspects of spoken words. Different representational levels accommodate phonetic variation in different ways, allowing listeners to overcome variability in the speech signal.

Representations are differently weighted depending on the specific demands of the linguistic situation. The specific context will drive listeners' attention to process the input in one way or the other. For implicit processing, where pronunciation does not need to be directly accessed, listeners accommodate deviations in the signal relatively easily. When more explicit processing requires making more specific judgments, short-term adjustment mechanisms will be applied. Focusing on pronunciation in detail increases the priority of phonetic-phonological representations. Group differences in sound discrimination ability at the phonetic level, such as those caused by difference in L1 or age of L2 acquisition, produce differences in explicit word recognition.

Our results make it clear that in the investigation of perception of dialectal variation, researchers will need to pay close attention to the linguistic profile of the listeners who are tested, and the nature of the tests themselves. Our findings demonstrate that the success of finely-tuned accommodation to dialectal variation in our listeners' second language depended on both the person's first language and the age at which Basque was learned. Moreover, not only does linguistic experience affect performance, but so does the attentional set listeners adopt to do the required tasks. Ultimately, these mechanisms are responsible for perceptual adjustment, and as a result, for successful communication.

Acknowledgments

This study was conducted with the support of the PSI 2010-17781 grant to the second author from the Spanish Government (MINECO). We are grateful to our participants for their kindness and cooperation. We are in debt to Doug Davidson for the invaluable advice he provided to us regarding the statistical analyses. We also very much appreciate the extremely constructive feedback we received during the reviewing process.

References

Baayen, H. (2012). Mixed-effects models. In A. C. Cohn, C. Fougeron & M. K. Huffman (Eds.), *The Oxford handbook of laboratory phonology* (pp. 668 – 677): Oxford: Oxford University Press.

Baayen, H., Davidson, D., & Bates, D. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*(4), 390-412.

Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (submitted). Parsimonious mixed models. *arXiv preprint arXiv:1506.04967*.

Best, C. T., McRoberts, G. W., & Goodell, E. (2001). Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listener's native phonological system. *The Journal of the Acoustical Society of America, 109*(2), 775.

Best, C. T., McRoberts, G. W., & Sithole, N. M. (1988). Examination of perceptual reorganization for nonnative speech contrasts: Zulu click discrimination by English-speaking adults and infants. *Journal of Experimental Psychology: Human Perception and Performance, 14*(3), 345.

Best, C. T., & Strange, W. (1992). Effects of phonological and phonetic factors on cross-language perception of approximants. *Journal of Phonetics, 20*(3), 305-330.

Boersma, P. (2002). Praat, a system for doing phonetics by computer. *Glot international, 5*(9/10), 341-345.

Bosch, L., Costa, A., & Sebastian-Galles, N. (2000). First and second language vowel perception in early bilinguals. *European Journal of Cognitive Psychology, 12*(2), 189-221.

Broersma, M., & Cutler, A. (2008). Phantom word activation in L2. *System, 36*(1), 22-34.

Caramazza, A., Yeni-Komshian, G. H., Zurif, E. B., & Carbone, E. (1973). The acquisition of a new phonological contrast: The case of stop consonants in French-English bilinguals. *The Journal of the Acoustical Society of America, 54*(2), 421-428.

Catford, J. C. (1977). *Fundamental problems in phonetics* (Vol. 1): Edinburgh University Press Edinburgh.

Clarke, C. M., & Garrett, M. F. (2004). Rapid adaptation to foreign-accented English. *The Journal of the Acoustical Society of America, 116*(6), 3647-3658.

Connine, C. M., Blasko, D. G., & Titone, D. (1993). Do the beginnings of spoken words have a special status in auditory word recognition? *Journal of Memory and Language, 32*(2), 193-210.

Cutler, A. (2012). *Native listening: Language experience and the recognition of spoken words*: MIT Press.

Diaz, B., Mitterer, H., Broersma, M., & Sebastian-Galles, N. (2012). Individual differences in late bilinguals' L2 phonological processes: From acoustic-phonetic analysis to lexical access. *Learning and Individual Differences, 22*(6), 680-689.

Dixon, P. (2008). Models of accuracy in repeated-measures designs. *Journal of Memory and Language, 59*(4), 447-456.

Donegan, P., & Stampe, D. (2009). Hypotheses of Natural Phonology. *Poznan Studies in Contemporary Linguistics, 45 (1)*, 1-31.

Dufour, S., Kriegel, S., Alleesaib, M., & Nguyen, N. (2014). The perception of the French /s/-/ʃ/ contrast in early Creole-French bilinguals. *Language Sciences, 5*, 1-8.

Duñabeitia, J. A., Cholin, J., Corral, J., Perea, M., & Carreiras, M. (2010). SYLLABARIUM: An online application for deriving complete statistics for Basque and Spanish orthographic syllables. *Behavior research methods, 42*(1), 118-125.

Escudero, P., & Boersma, P. (2004). Bridging the gap between L2 speech perception research and phonological theory. *Studies in Second Language Acquisition, 26*(4), 551-585.

Flege, J. E. (1999). Age of learning and second language speech. *Second language acquisition and the critical period hypothesis*, 101-131.

Flege, J. E., & MacKay, I. R. A. (2004). Perceiving vowels in a second language. *Studies in Second Language Acquisition, 26*(1), 1-34.
Floccia, C., Goslin, J., Girard, F., & Konopczynski, G. (2006). Does a regional accent perturb speech processing? *Journal of Experimental Psychology: Human Perception and Performance, 32*(5), 1276-1293.
Garcia Lecumberri, M. L., Cooke, M., & Cutler, A. (2010). Non-native speech perception in adverse conditions: A review. *Speech Communication, 52*(11-12), 864-886.
Gaskell, M. G., & Marslen-Wilson, W. D. (1997). Integrating form and meaning: A distributed model of speech perception. *Language and Cognitive Processes, 12*(5-6), 613-656.
Hualde, J. I. (2010). Neutralización de sibilantes vascas y seseo en castellano. *Oihenart: cuadernos de lengua y literatura*(25), 89-116.
Hualde, J. I., & Ortiz de Urbina, J. (2003). *A grammar of Basque* (Vol. 26): Walter de Gruyter.
Iverson, P., Kuhl, P. K., Akahane-Yamada, R., Diesch, E., Tohkura, Y., Kettermann, A., & Siebert, C. (2003). A perceptual interference account of acquisition difficulties for non-native phonemes. *Cognition, 87*(1), B47-B57.
Johnson, J. S., & Newport, E. L. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology, 21*(1), 60-99.
Kim, K., Relkin, N. R., Lee, K. M., & Hirsch, J. (1997). Distinct cortical areas associated with native and second languages. *Nature, 388*(6638), 171-174.
Kuhl, P. K. (1994). Learning and representation in speech and language. *Current Opinion in Neurobiology, 4*(6), 812-822.
Ladefoged, P., & Maddieson, I. (1996). *The sounds of the world's languages*: Blackwell Oxford.
LoCasto, P. C., & Connine, C. M. (2002). Rule-governed missing information in spoken word recognition: Schwa vowel deletion. *Perception & Psychophysics, 64*(2), 208-219.
MacLeod, B. (2012). *The Effect of Perceptual Salience on Phonetic Accommodation in Cross-Dialectal Conversation in Spanish.* University of Toronto.
Marslen-Wilson, W. D., & Zwitserlood, P. (1989). Accessing spoken words: The importance of word onsets. *Journal of Experimental Psychology: Human Perception and Performance, 15*(3), 576-585.
McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology, 18*(1), 1-86.
McQueen, J. M., & Cutler, A. (2001). Spoken word access processes: An introduction. *Language and Cognitive Processes, 16*(5-6), 469-490.
Meyer, D. E., & Schvaneveldt, R. W. (1971). Facilitation in recognizing pairs of words: evidence of a dependence between retrieval operations. *Journal of Experimental Psychology, 90*(2), 227-234.
Mitterer, H., & Tuinman, A. (2012). The role of native-language knowledge in the perception of casual speech in a second language. *Frontiers in Psychology, 3*, 1-13.
Navarra, J., Sebastian-Galles, N., & Soto-Faraco, S. (2005). The perception of second language sounds in early bilinguals: new evidence from an implicit measure. *Journal of Experimental Psychology: Human Perception and Performance, 31*(5), 912-918.
Norris, D. (1994). Shortlist: A connectionist model of continuous speech recognition. *Cognition, 52*(3), 189-234.
Norris, D., McQueen, J. M, & Cutler, A. (2000). Merging information in speech recognition: Feedback is never necessary. *Behavioral and Brain Sciences, 23*(03), 299-325.
Norris, D., & McQueen, J. M. (2008). Shortlist B: a Bayesian model of continuous speech recognition. *Psychological Review, 115*(2), 357-395.

Norris, D., McQueen, J. M., & Cutler, A. (2003). Perceptual learning in speech. *Cognitive Psychology, 47*(2), 204-238.

Oñederra, M. L. (2009). Early bilingualism as a source of morphonological rules for the adaptation of loanwords: Spanish loanwords in Basque. *A. Calabrese & WL Wetzels (arg.), Loan Phonology, John Benjamins, Amsterdam & Philadelphia*, 193-210.

Pallier, C., Bosch, L., & Sebastian-Galles, N. (1997). A limit on behavioral plasticity in speech perception. *Cognition, 64*(3), B9-B17.

Pallier, C., Colomé, A., & Sebastian-Galles, N. (2001). The influence of native-language phonology on lexical access: Exemplar-based versus abstract lexical entries. *Psychological Science, 12*(6), 445-449.

Pisoni, D. B., & Luce, P. A. (1987). Acoustic-phonetic representations in word recognition. *Cognition, 25*(1), 21-52.

Pitt, M. A. (2009). How are pronunciation variants of spoken words recognized? A test of generalization to newly learned words. *Journal of Memory and Language, 61*(1), 19-36.

Pitt, M. A., & Szostak, C. M. (2012). A lexically biased attentional set compensates for variable speech quality caused by pronunciation variation. *Language and Cognitive Processes, 27*(7-8), 1225-1239.

Polka, L. (1995). Linguistic influences in adult perception of non-native vowel contrasts. *The Journal of the Acoustical Society of America, 97*, 1286-1296.

Polka, L., Colantonio, C., & Sundara, M. (2001). A cross-language comparison of /d/–/ð/ perception: Evidence for a new developmental pattern. *The Journal of the Acoustical Society of America, 109*(5), 2190-2201.

Ranbom, L., & Connine, C. M. (2007). Lexical representation of phonological variation in spoken word recognition. *Journal of Memory and Language, 57*(2), 273-298.

Samuel, A. G., & Kraljic, T. (2009). Perceptual learning for speech. *Attention, Perception, & Psychophysics, 71*(6), 1207-1218.

Samuel, A. G., & Larraza, S. (2015). Does listening to non-native speech impair speech perception? *Journal of Memory and Language, 81*, 51-71.

Sebastian-Galles, N., & Baus, C. (2005). On the relationship between perception and production in L2 categories. In Anne Cutler (Ed.), *Twenty-first century psycholinguistics: Four cornerstones* (pp. 279-292): Psychology Press.

Sebastian-Galles, N., Echeverría, S., & Bosch, L. (2005). The influence of initial exposure on lexical representation: Comparing early and simultaneous bilinguals. *Journal of Memory and Language, 52*(2), 240-255.

Sebastian-Galles, N., Rodríguez-Fornells, A., de Diego-Balaguer, R., & Diaz, B. (2006). First-and second-language phonological representations in the mental lexicon. *Journal of Cognitive Neuroscience, 18*(8), 1277-1291.

Sebastian-Galles, N., & Soto-Faraco, S. (1999). Online processing of native and non-native phonemic contrasts in early bilinguals. *Cognition, 72*(2), 111-123.

Silverberg, S., & Samuel, A. G. (2004). The effect of age of second language acquisition on the representation and processing of second language words. *Journal of Memory and Language, 51*(3), 381-398.

Sumner, M. (2011). The role of variation in the perception of accented speech. *Cognition, 119*(1), 131-136.

Sumner, M., & Samuel, A. G. (2005). Perception and representation of regular variation: The case of final /t/. *Journal of Memory and Language, 52*(3), 322-338.

Sumner, M., & Samuel, A. G. (2009). The effect of experience on the perception and representation of dialect variants. *Journal of Memory and Language, 60*(4), 487-501.

Sundara, M., & Polka, L. (2008). Discrimination of coronal stops by bilingual adults: The timing and nature of language interaction. *Cognition, 106*(1), 234-258.
Tuinman, A., Mitterer, H., & Cutler, A. (2012). Resolving ambiguity in familiar and unfamiliar casual speech. *Journal of Memory and Language, 66*(4), 530-544.
Txillardegi. (1980). *Euskal Fonologia* (Ediciones Vascas Ed.). San Sebastian.
Weber, A., & Cutler, A. (2004). Lexical competition in non-native spoken-word recognition. *Journal of Memory and Language, 50*(1), 1-25.
Witteman, M. J., Weber, A., & McQueen, J. M. (2013). Foreign accent strength and listener familiarity with an accent codetermine speed of perceptual adaptation. *Attention, Perception, & Psychophysics, 75*(3), 537-556.
Zawiszewski, A., Gutiérrez, E., Fernández, B., & Laka, I. (2011). Language distance and non-native syntactic processing: Evidence from event-related potentials. *Bilingualism: Language and Cognition, 14*(3), 400-411.
Zhang, X., & Samuel, A. G. (2015). The activation of embedded words in spoken word recognition. *Journal of Memory and Language, 79*, 53-75.

Appendix A

Summary of the models built for each experiment and associated lmer syntax

	lmer model syntax*
Exp1	`L1 * AOA * cond + position + (1+cond|subject) + (1+L1|item)`
Exp2	`L1 * AOA * cond * sound + (1+cond|subject) + (1+L1|item)`
Exp3	`L1 * AOA * cond + context + (1+cond|subject) + (1+L1|item)`

*Treatment coding was used in the models.

Appendix B

Stimuli used in the Auditory Lexical Decision Task

Z-words	S-nonwords	S-words	Z-nonwords	Control Words	Control Nonwords	Control Nonwords
ugazaba	ugasaba	egunpasa	egunpaza	aldaketa	aperatsa	tilista
azalora	asalora	lasaigarria	lazaigarria	arrantza	apiadura	izenburrua
diruzaina	dirusaina	osaba	ozaba	bazterra	barkanena	adinena
auzolana	ausolana	eurijasa	eurijaza	bilaketa	bikodea	adiskitea
izeba	iseba	basamortua	bazamortua	egoera	bitea	gokorra
oinazea	oinasea	hamasei	hamazei	egoitza	buruzakia	iraubena
adierazi	adierasi	ausarta	auzarta	eguna	eurria	aldaga
erizaina	erisaina	asaldatu	azaldatu	elurra	garbikeda	batasuma
ehiza	ehisa	besapea	bezapea	erabilera	guta	begirata
ikustezina	ikustesina	basajauna	bazajauna	euskara	harama	naskorra
lizardia	lisardia	erasoa	erazoa	gaitza	helburrua	perdina
buruzagia	burusagia	isilunea	izilunea	herrialdea	hetadura	deikarria
nekazaria	nekasaria	esaera	ezaera	lasterketa	igastaroa	erapakia
gazia	gasia	babesa	babeza	maila	ipilera	ekinena
gizartea	gisartea	desagertu	dezagertu	oharra	jadorria	txorinaloa
azeria	aseria	osaera	ozaera	oinarria	lekea	lebokoa
noraeza	noraesa	jasangarria	jazangarria	orbela	meska	eskularua
luzera	lusera	usaina	uzaina	ordua	ondasuma	lapana
bezeroa	beseroa	haserrea	hazerrea	oreina	orerkia	nailua
geruza	gerusa	lehengusua	lehenguzua	sendia	pizkarra	argindara
azalera	asalera	esaldia	ezaldia	urratsa	sakarra	txinista
atezaina	atesaina	gaupasa	gaupaza	urtea	soinegoa	simismena
gozamena	gosamena	erosoa	erozoa	laguntza	tragetsa	asperdurra
ezaguna	esaguna	artisautza	artizautza	zorrotza	txostema	nalkoa

Appendix C

The speaker who recorded stimuli of Experiment 2 was not a Western dialect speaker himself, but a speaker with the /s̻/-/s̺/ contrast who normally pronounces these distinctively. So, one might wonder to what extent the /s̺/ > /s̻/ mispronunciations by him sound like the neutralizations of a native Western speaker. To measure this, we extracted the apical sounds from the s-nonwords he produced and compared them with apical sounds pronounced by a native Western speaker producing the same nonwords. This analysis showed significant differences in spectral centroid ($t(8)=29.7$, $p<.001$) and duration ($t(8)=3.06$, $p<.05$). This suggests that the neutralized fricative by a Western speaker is not acoustically the same as the apical fricative by a non-neutralizing Basque native speaker. Note that these are just two speakers, so it is difficult to know whether these are individual differences, or more general ones. If listeners were able to perceive these differences, S-nonwords we present in the task should be rejected more often and thus, 'Nonword' responses would dominate. This would mean that the acoustic differences of S-nonwords, compared to a Western speaker's pronunciation, could hinder listeners' ability to adjust to dialectal variation.

Listeners might also expect a clear and systematic distinction of Basque fricatives by someone who has shown that he can differentiate them. Thus, the design and context of the lexical decision task we present here is done in such a way that it may increase 'Nonword' responses for S-nonwords by participants.

Appendix D

A Representative Sample of the Stimuli Sets Used in Auditory Semantic Priming

	CONDITION	PRIME	TARGET
Trial 1	Related	**zilarra**	**urrea**
	Unrelated	**kiroldegia**	**begia**
	Control Nonword	**gezura**	**egia**
	Accented variant	**serua**	**urdina**
Trial 2	Related	**bizarra**	**bibotea**
	Unrelated	**galdetu**	**emakumea**
	Control Nonword	**pizitza**	**heriotza**
	Accented variant	**etorkisuna**	**oraina**
Trial 3	Related	**miazkatu**	**mingaina**
	Unrelated	**organoa**	**gaua**
	Control Nonword	**azara**	**leuna**
	Accented variant	**irabasi**	**galdu**
Trial 4	Related	**izua**	**beldurra**
	Unrelated	**euskalkia**	**ilea**
	Control Nonword	**zigina**	**garbia**
	Accented variant	**erastuna**	**behatza**
Trial 5	Related	**izena**	**deitura**
	Unrelated	**ibilkera**	**odola**
	Control Nonword	**zurra**	**egurra**
	Accented variant	**suhaitza**	**hostoa**
Trial 6	Related	**zezena**	**adarra**
	Unrelated	**eskubidea**	**okerra**
	Control Nonword	**bozik**	**alai**
	Accented variant	**isotza**	**hotza**
Trial 7	Related	**eguzkia**	**uda**
	Unrelated	**pirata**	**berdea**
	Control Nonword	**zaurrixa**	**mina**
	Accented variant	**sapixa**	**mukiak**
Trial 8	Related	**zenbakia**	**bat**
	Unrelated	**mila**	**eskola**
	Control Nonword	**ziraria**	**alaitasuna**
	Accented variant	**sergia**	**ordaindu**
Trial 9	Related	**zozketa**	**loteria**
	Unrelated	**izaera**	**konponbidea**
	Control Nonword	**narrazkixa**	**margoa**
	Accented variant	**subixa**	**ibaia**

Trial 10	Related	**zaldia**	**behorra**
	Unrelated	**papergintza**	**untxia**
	Control Nonword	**oztia**	**garia**
	Accented variant	**asokia**	**fruta**

Figure captions

FIGURE 1

Distribution of the apical sound in Spanish and Basque based on its spectral centroid and peak amplitude values.

FIGURE 2

Accuracy and RT results for the AXB task by the different groups of Basque bilinguals. Results for AAB and ABB positions were merged in this Figure. Error bars represent standard errors.

FIGURE 3

Accuracy and RT results for words in the LDT task by the different groups of Basque bilinguals. Error bars represent standard errors.

FIGURE 4

Acceptance rates and RT results for nonwords in the LDT task by the different groups of Basque bilinguals. Note that the left panel shows the percentage of nonwords that were taken as valid Basque utterances, showing that most of the time all five groups accepted S-nonwords. Results plotted in the right panel correspond to rejected nonwords. Error bars represent standard errors.

FIGURE 5

Reaction time results for the different stimulus types used in the LDT task. The two comparisons of interest are included: Z-words versus S-nonwords, and S-nonwords versus Z-nonwords, in order to see how L2 licensed and unlicensed dialectal variation is accommodated. Recall that results plotted in the 'S-nonword' and 'Z-nonword' bars correspond to nonwords that were accepted as valid Basque utterances, complementary to the reaction times of rejected nonwords plotted in Figure 4 (right panel). Error bars represent standard errors.

FIGURE 6

Correlation between the accuracy on the critical contrast discrimination (AXB task) and the acceptance of Z-nonwords (LDT task), by the different groups of bilinguals. Error bars represent standard errors.

FIGURE 7

Semantic priming effects, based on the reaction times, for the five participant groups across the four experimental conditions. Error bars represent standard errors.

Table captions

TABLE 1

Phonemic distribution of the apico- and lamino-alveolar voiceless fricatives across the three languages of interest (i.e., Basque, Spanish, French), taking Basque dialects into account. Corresponding graphemes for the Basque sounds are given in quotations.

TABLE 2

Acoustic characteristics of the Basque **/s̻/** and **/s̺/** fricatives extracted from nonwords used in the AXB task.

TABLE 3

Acoustic characteristics of the Basque **/s̻/** and **/s̺/** fricatives extracted from words and nonwords used in the LDT task.

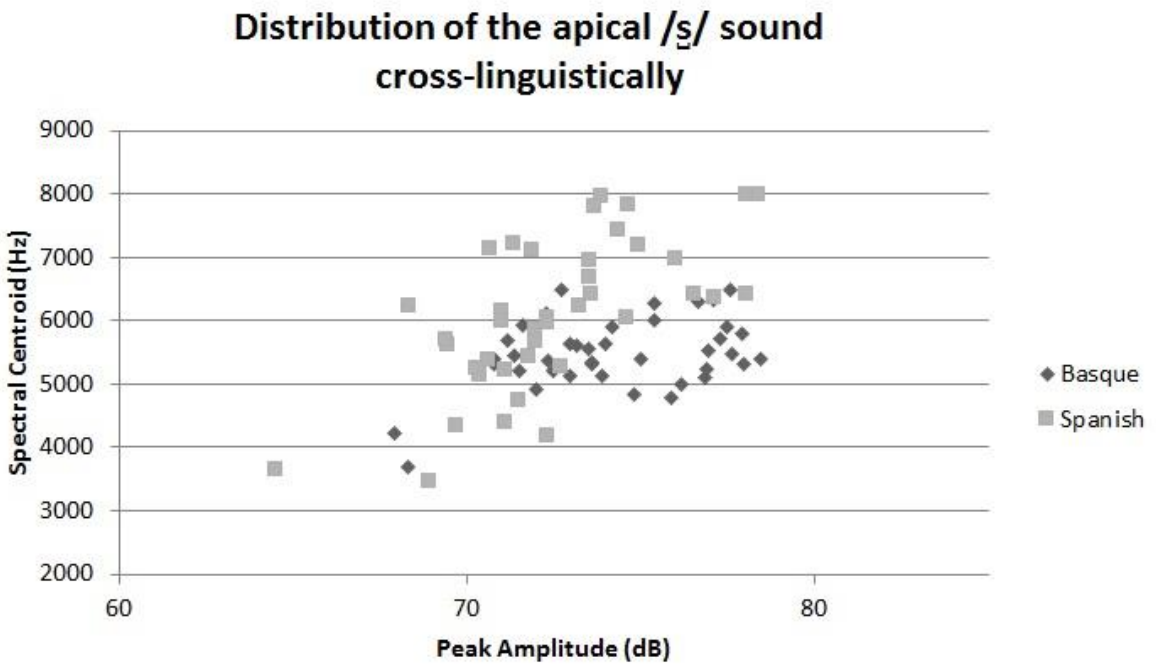
TABLE 4

Examples of single-feature-change and 2-feature-change trials used in the Auditory Semantic Priming experiment. Phonetic features that change in Accented and Control conditions relative to the Related condition are in bold.

TABLE 5

Accuracy and RT results for the auditory semantic priming task by Basque bilinguals with different linguistic profiles.

Figure
Click here to download Figure AccentedSpeech_Figures&Tables_final.docx



	Basque /s̻/	Spanish /s̻/
Peak Amplitude (dB)	M 74.2 (sd 2.7)	M 72.5 (sd 2.9)
Spectral Centroid (Hz)	M 5478 (sd 567.3)	M 6084 (sd 1178.6)
Duration (s)	M 0.15 (sd 0.02)	M 0.14 (sd 0.01)

FIGURE 1

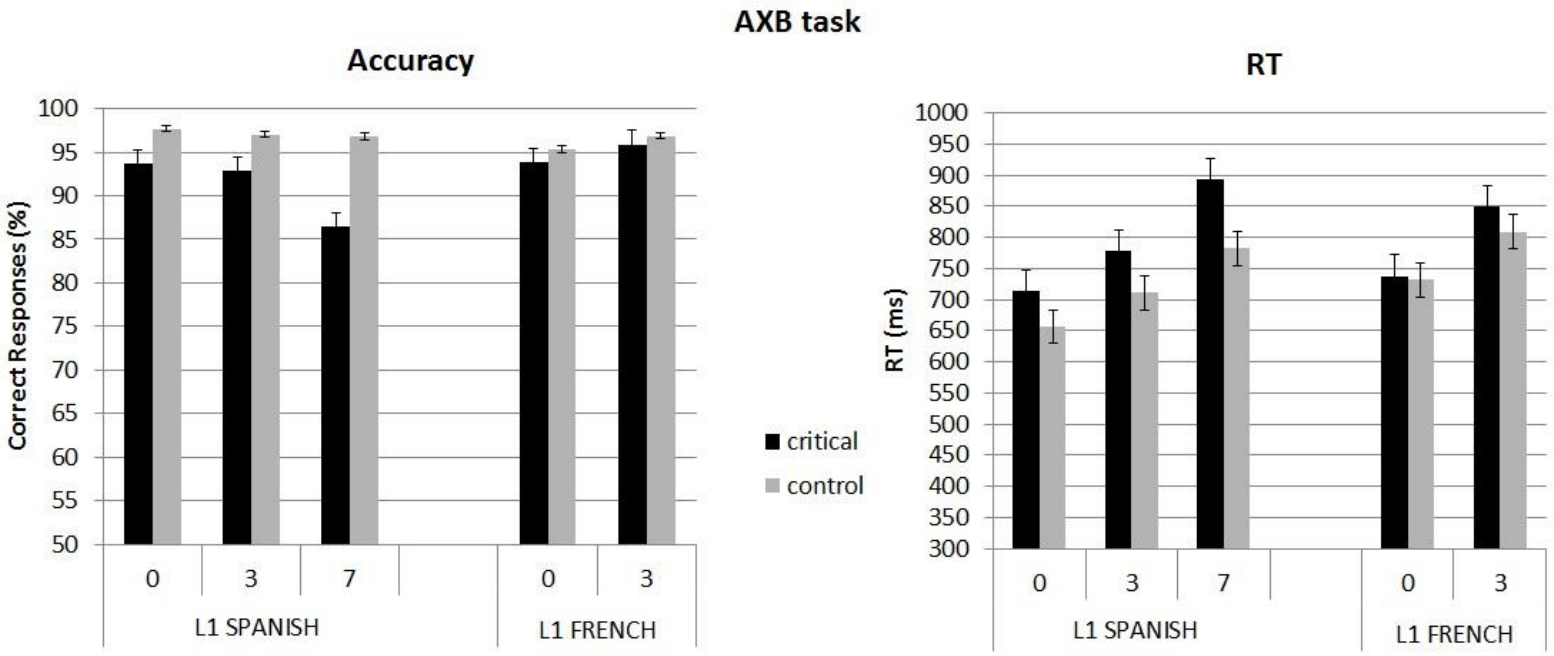

AXB task
Accuracy
Correct Responses (%)
100
95
90
85
80
75
70
65
60
55
50
0
3
7
0
3
L1 SPANISH
L1 FRENCH
critical
control
RT
RT (ms)
1000
950
900
850
800
750
700
650
600
550
500
450
400
350
300
0
3
7
0
3
L1 SPANISH
L1 FRENCH

FIGURE 2

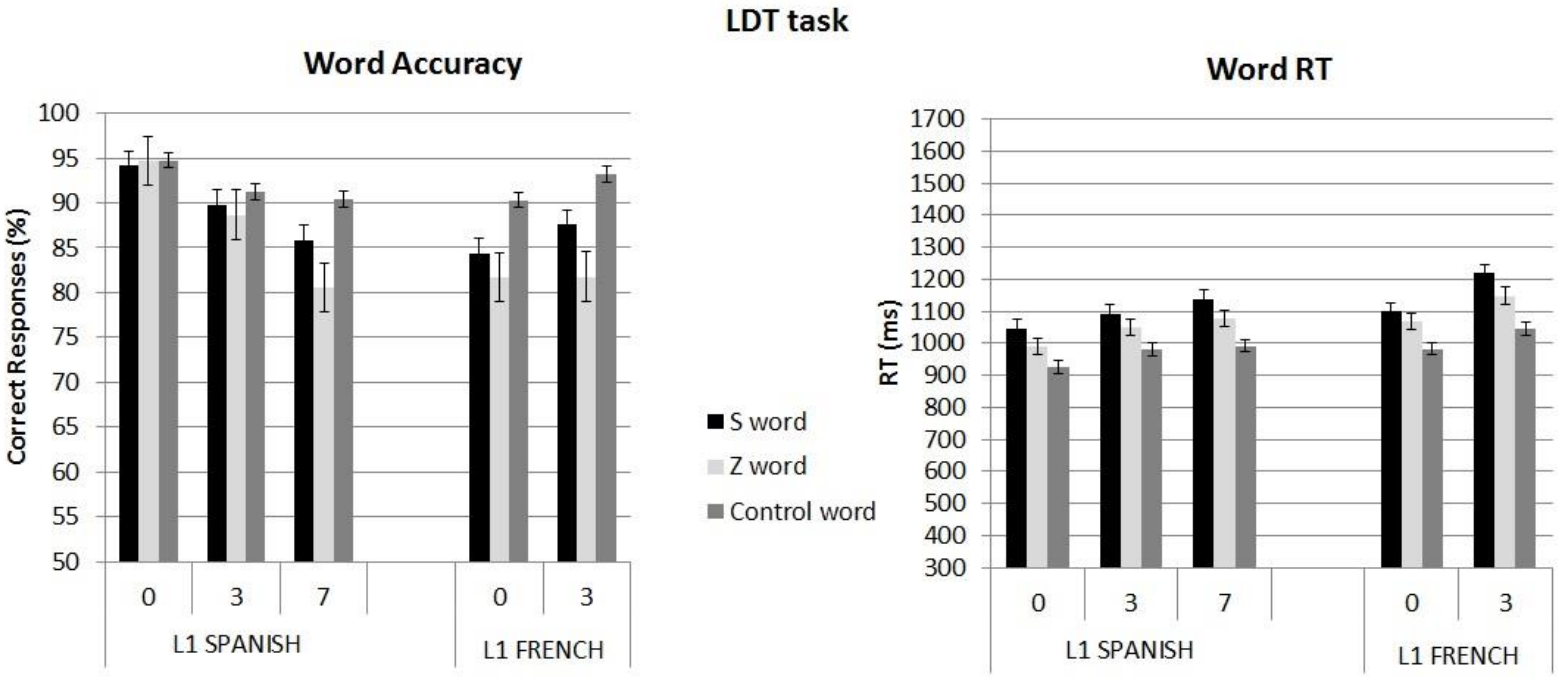

LDT task
Word Accuracy
Correct Responses (%)
100
95
90
85
80
75
70
65
60
55
50
0
3
7
0
3
L1 SPANISH
L1 FRENCH
S word
Z word
Control word
Word RT
RT (ms)
1700
1600
1500
1400
1300
1200
1100
1000
900
800
700
600
500
400
300
0
3
7
0
3
L1 SPANISH
L1 FRENCH

FIGURE 3

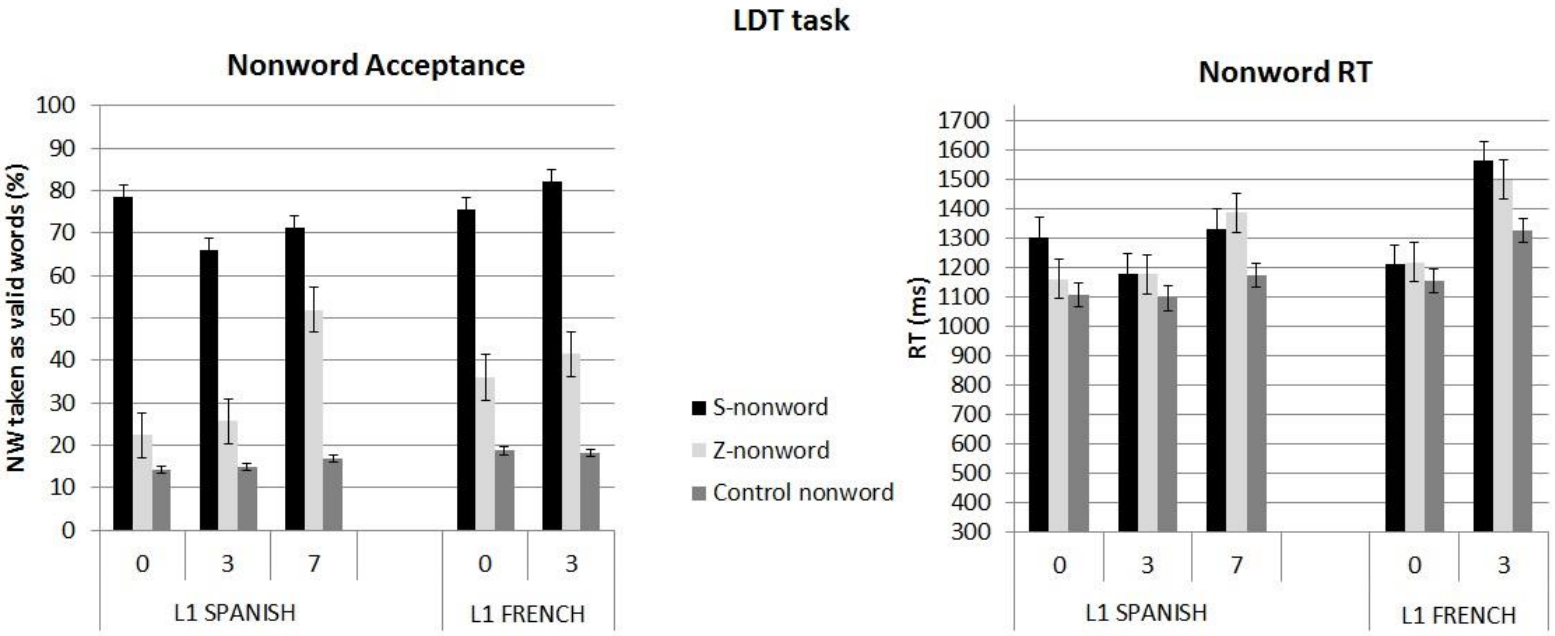

LDT task
Nonword Acceptance
NW taken as valid words (%)
100
90
80
70
60
50
40
30
20
10
0
0
3
7
0
3
L1 SPANISH
L1 FRENCH
S-nonword
Z-nonword
Control nonword
Nonword RT
RT (ms)
1700
1600
1500
1400
1300
1200
1100
1000
900
800
700
600
500
400
300
0
3
7
0
3
L1 SPANISH
L1 FRENCH

FIGURE 4

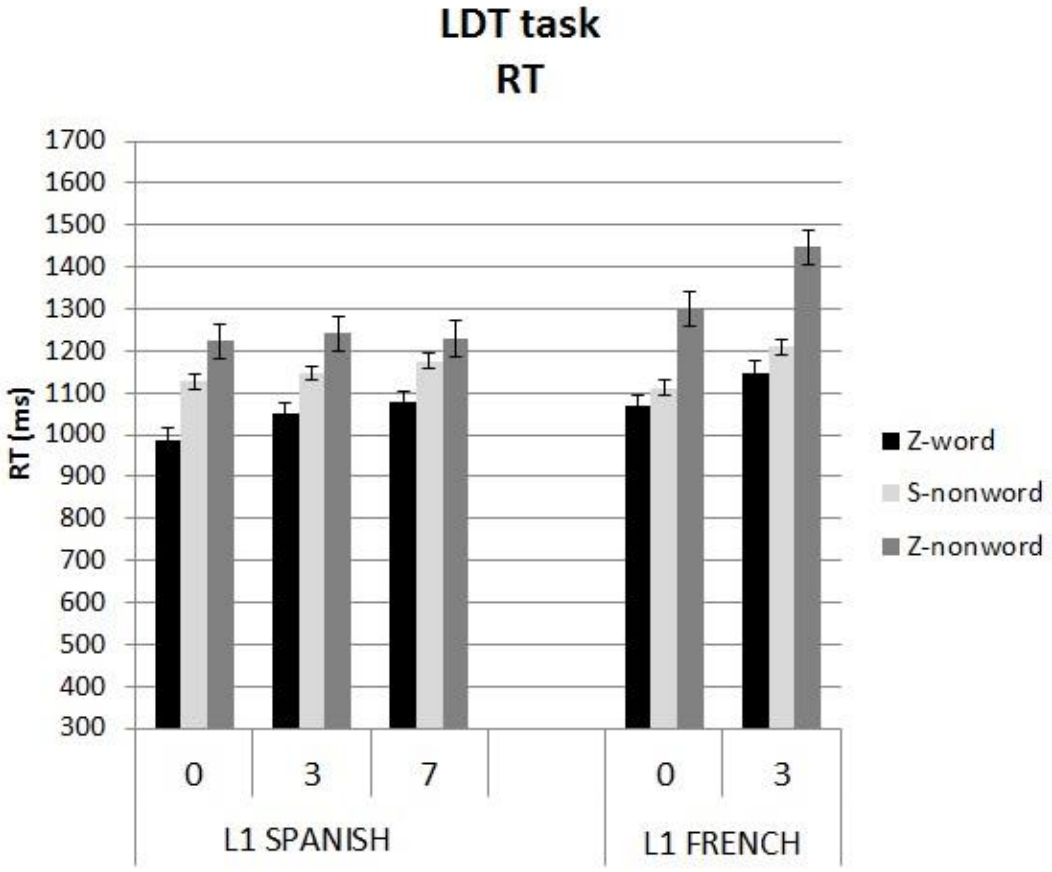

LDT task
RT
RT (ms)
1700
1600
1500
1400
1300
1200
1100
1000
900
800
700
600
500
400
300
0
3
7
0
3
L1 SPANISH
L1 FRENCH
Z-word
S-nonword
Z-nonword

FIGURE 5

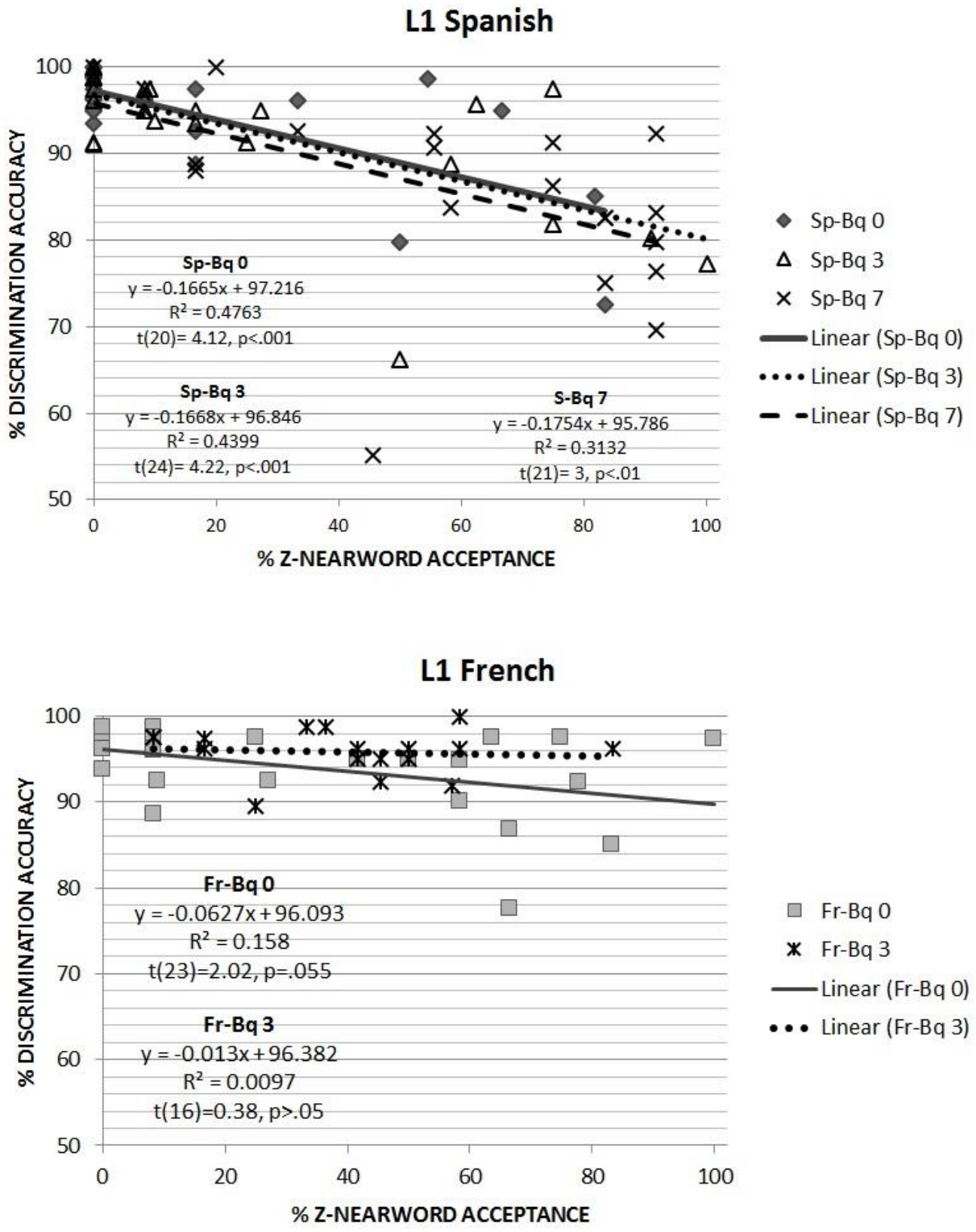

L1 Spanish
% DISCRIMINATION ACCURACY
% Z-NEARWORD ACCEPTANCE
Sp-Bq 0
y = -0.1665x + 97.216
R² = 0.4763
t(20)= 4.12, p<.001
Sp-Bq 3
y = -0.1668x + 96.846
R² = 0.4399
t(24)= 4.22, p<.001
S-Bq 7
y = -0.1754x + 95.786
R² = 0.3132
t(21)= 3, p<.01
Sp-Bq 0
Sp-Bq 3
Sp-Bq 7
Linear (Sp-Bq 0)
Linear (Sp-Bq 3)
Linear (Sp-Bq 7)
L1 French
% DISCRIMINATION ACCURACY
% Z-NEARWORD ACCEPTANCE
Fr-Bq 0
y = -0.0627x + 96.093
R² = 0.158
t(23)=2.02, p=.055
Fr-Bq 3
y = -0.013x + 96.382
R² = 0.0097
t(16)=0.38, p>.05
Fr-Bq 0
Fr-Bq 3
Linear (Fr-Bq 0)
Linear (Fr-Bq 3)

FIGURE 6

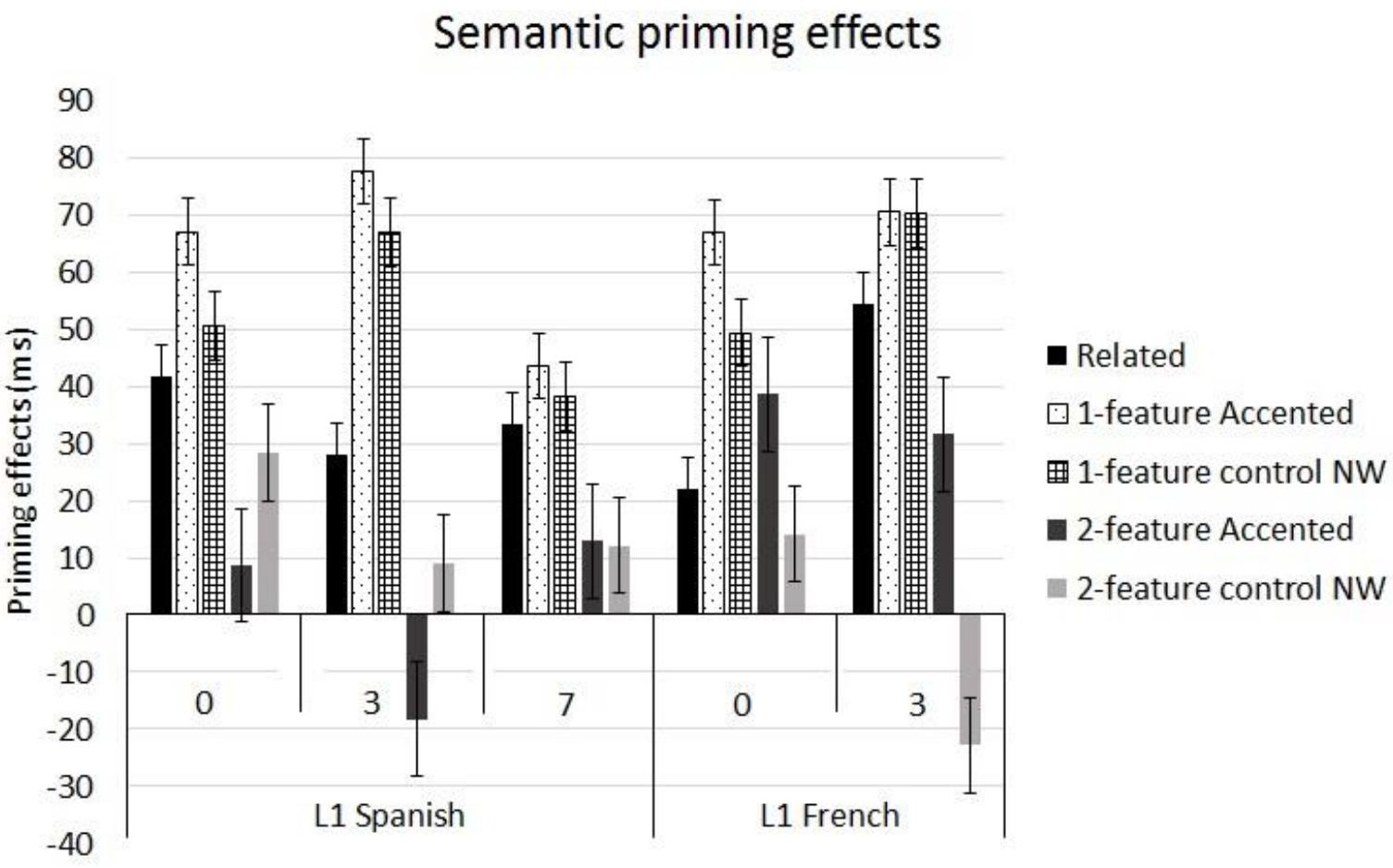

Semantic priming effects
Priming effects (ms)
90
80
70
60
50
40
30
20
10
0
-10
-20
-30
-40
0
3
7
0
3
L1 Spanish
L1 French
Related
1-feature Accented
1-feature control NW
2-feature Accented
2-feature control NW

FIGURE 7

	Basque		Castilian Spanish	French
	Standard dialect	**Western dialect**		
Apico-alveolar fricative	**/s̻/ “S”**	**/s̻/**	**/s̻/**	**-**
Lamino-alveolar fricative	**/s̺/ “Z”**		**-**	**/s̺/**

TABLE 1

	S apical M(sd)	Z laminal M(sd)	t – value *p* value
Peak Amplitude (dB)	69.8 (4.77)	62.7 (5.63)	$t(79)=9.52$ $p<.001$
Spectral Centroid (Hz)	5292 (638.48)	9437 (1554.15)	$t(79)=26.63$ $p<.001$
Duration (s)	0.16 (0.018)	0.169 (0.026)	$t(79)=3.55$ $p<.001$

TABLE 2

	S-words vs Z-nonwords			Z-words vs S-nonwords		
	S-words M(sd)	**Z-nonwords** M(sd)	t – value *p* value	**Z-words** M(sd)	**S-nonwords** M(sd)	t – value *p* value
Peak amplitude (dB)	71.4 (1.97)	63.8 (1.94)	*t*(23)=23.37 *p*<.001	61.8 (2.52)	71.2 (2.42)	*t*(23)=21.62 *p*<.001
Spectral centroid (Hz)	5896 (353.24)	7331 (826.99)	*t*(23)=10.17 *p*<.001	7221 (902.14)	5854 (278.27)	*t*(23)=8.71 *p*<.001
Duration (s)	0.115 (0.011)	0.120 (0.0118)	*t*(23)=1.34 *p*>.05	0.118 (0.0134)	0.122 (0.0121)	*t*(23)=1.1 *p*>.05

TABLE 3

CONDITION	*single-feature-change*		*2-feature-change*	
	PRIME	TARGET	PRIME	TARGET
1. Related prime	*betazala* 'the eyelid'	*begia* 'the eye'	*emaztea* 'the wife'	*senarra* 'the husband'
2. Unrelated prime	*kiroldegia* 'the sports center'		*galanta* 'the great one'	
3. Accented prime (Western variant)	**beta**s**ala*		**ema**st**i**a***	
4. Control Nonword	**be**d**azala*		**e**n**azti**a***	

TABLE 4

Semantic Priming task

Accuracy (correct %)						
	L1 Spanish			L1 French		MEAN (sd)
Condition	0	3	7	0	3	
Related	98.8	98.1	94.7	97.7	98.3	**97.5 (1.6)**
Unrelated	96.7	93.2	94.3	94.2	96.6	**95 (1.6)**
1-feature Accented	100	98.7	97.7	98.7	99.1	**98.8 (0.8)**
1-feature Control Nonword	100	96.8	100	97.4	98.3	**98.5 (1.5)**
2-feature Accented	99.4	97.4	93.9	94.2	94.1	**95.8 (2.5)**
2-feature Control Nonword	98.8	96.8	93.1	97.4	97.5	**96.7 (2.1)**
MEAN (sd)	**98.9 (1.2)**	**96.8 (1.9)**	**95.6 (2.7)**	**96.6 (1.9)**	**97.3 (1.8)**	

RT (ms)						
	L1 Spanish			L1 French		MEAN (sd)
Condition	0	3	7	0	3	
Related	873	905	963	935	945	**924 (35.3)**
Unrelated	915	933	996	957	999	**960 (37.4)**
1-feature Accented	848	856	953	891	929	**895 (45.4)**
1-feature Control Nonword	864	867	958	908	929	**905 (40.3)**
2-feature Accented	906	952	983	919	967	**945 (32.4)**
2-feature Control Nonword	886	924	984	943	1022	**952 (52.5)**
MEAN (sd)	**882 (25.5)**	**906 (38.1)**	**973 (17.3)**	**926 (24.5)**	**965 (38.6)**	

TABLE 5