This is a contribution from *Romance Languages and Linguistic Theory 2012. Selected papers from 'Going Romance' Leuven 2012*.
Edited by Karen Lahousse and Stefania Marzo.

A comparison of fricative voicing and lateral velarization phenomena in Barcelona

A variationist approach to Spanish in contact with Catalan

Justin Davidson
University of Illinois at Urbana-Champaign

This investigation constitutes a quantitative variationist approach toward Spanish in contact with Catalan in Barcelona, Spain. It seeks to empirically measure concrete usage patterns of two phonetic variants, [ɬ] and [z], in the Spanish of Catalan-Spanish bilinguals, as well as establish the extent to which both variants are conditioned by linguistic factors and Catalan dominance. The careful Spanish speech of 20 Barcelonan females (ages 18–27) was elicited through a word-reading task. Goldvarb binomial logistic regression analyses revealed that sensitivity to linguistic factors varied according to Catalan dominance. Moreover, although both variants were favored most by Catalan-dominant speakers, usage patterns among more Spanish-dominant speakers were divergent, consistent with claims of negative social value linked solely to [ɬ].

1. Introduction[1]

This study examines language use patterns of two features of Spanish in contact with Catalan, that is, the Spanish spoken by bilingual Catalan-Spanish speakers

1. This research was generously funded by the Tinker Foundation (through the UIUC Center for Latin American & Caribbean Studies) and the Anthony M. Pasquariello Research Award. I extend my sincere thanks to Antonio Torres Torres, Ana María Fernández Planas, Pilar Ocaña, Joan Carles Bonilla Mora (*Universitat de Barcelona*), Carles Salse Capdevila (*Universitat Autònoma de Barcelona*), Mireia Trenchs, Pilar Prieto (*Universitat Pompeu Fabra*), Michael Newman (Queens College/CUNY), John Wilcox, and Jonathan MacDonald (UIUC), who facilitated my successful data collection in Barcelona. I am additionally indebted to the helpful comments and support from Anna María Escobar, Zsuzsanna Fagyal, and José Ignacio Hualde (UIUC), and lastly must thank Naomi Nagy (University of Toronto) for extensive consultations for statistical analyses. All remaining errors are my own. Correspondences welcome: jdavids2@illinois.edu.

in Barcelona, Spain. Spanish in contact with Catalan, henceforth referred to as Catalan Contact Spanish (CCS), is a Catalanized variety of Spanish that exhibits features of Catalan syntax, morphology, lexicon, and phonetics. We focus on usage patterns of two CCS phonetic phenomena: (1) velarization of the Spanish voiced alveolar lateral [l] to velarized [ɬ] (e.g. *al parque* [a**l**.páɾ.ke] > [aɬ.páɾ.ke] 'to the park'), and (2) voicing of the Spanish intervocalic voiceless alveolar fricative [s] to voiced [z] (e.g. *los años* [lo.sá.ɲos] > [lo.**z**á.ɲos] 'the years'). Their selection is motivated by their presence in previous treatments of Barcelonan CCS (cf. Sinner 2002; Vann 2001; Wesch 1997). By examining [ɬ] and [z] usage in a select group of Barcelonan CCS-speakers, this study seeks to establish, as conditioned by Catalan usage and exposure, both: (1) the degree to which each variant is present in formal CCS production and (2) which linguistic (or language-internal) factors condition each variant's production.

2. Review of CCS lateral velarization and intervocalic fricative voicing phenomena

2.1 Linguistic characterizations of lateral velarization and intervocalic fricative voicing

CCS lateral velarization and intervocalic fricative voicing have both been ascribed to influence from Catalan on the basis of substantial acoustic and articulatory research. Regarding the production of laterals, monolingual Spanish varieties are characterized as exhibiting a non-velarized lateral category (termed "clear" or "light") in all linguistic contexts, whereas Catalan is conversely characterized as featuring a velarized lateral category (termed "dark") in all linguistic contexts (Hofwegen 2009; 306; Hualde 2005; 179; Prieto 2004; 204; Recasens 1991, 305–306; 2004, 594; Recasens, Fontdevila & Pallarès 1995, 38). While the two lateral categories share a coronal place of articulation, the dark lateral is distinguished by an additional dorsal approximation at the velar region accompanied by pre-dorsal lowering (Simonet 2010, 664; and references within).

With respect to the production of (intervocalic) fricatives, both Spanish and Catalan feature an apical-alveolar voiceless /s/, articulated with a gesture of the tongue-tip toward the alveolar ridge (Hualde 2005; 47; Prieto 2004; 204–205; Quilis 1981; 234–235; Recasens 1991, 267–268). However, the majority of Catalan varieties (and for the purposes of this study, Central Catalan, spoken in Barcelona) feature an additional apical-alveolar phoneme, voiced /z/, which word-initially and word-medially contrasts phonemically with voiceless /s/ and produces minimal pairs such as *casa* /z/ 'house' and *caça* /s/ '3.s. hunts'. This phonemic voicing

contrast is neutralized word-finally, resulting in [s] or [z] depending on the following segment's voicing feature: when followed by a vowel, the word-final alveolar fricative is systematically voiced (e.g. *gos* [s] 'dog'; *gos estrany* [z] 'strange dog') by means of a phonological, post-lexical voicing rule that affects all Catalan word-final sibilants (e.g. /f/, /s/, /z/, /ʃ/, and /ʒ/) preceding a vowel-initial word (Bonet i Lloret 1998; 118–119; Pieras 1999; 212; Prieto 2004, 208, 216). Since alveolar /s/ is the only Catalan sibilant shared by Spanish (with a normal distribution, as Spanish /f/ is absent word-finally save in foreign borrowings [Hualde, Olarrea, Escobar & Travis 2010, 74]), the phenomenon of intervocalic fricative voicing used in this paper refers exclusively to alveolar fricatives. Thus, whereas voiced intervocalic alveolar fricatives are common in Catalan (resultant from /z/ and systematic voiced [z] in word-final intervocalic position), intervocalic [z] is unattested as a systematic feature of monolingual Spanish.

2.2 Past evidence of [ɬ] and [z] in Barcelonan CCS

Prior treatments of phonetic features of CCS discuss [ɬ] and [z] as products of Catalan-Spanish transfer that characterize a Catalanized variety of Spanish of Catalan-dominant speakers. For example, impressionistic interviews by Sinner (2002) suggest that [ɬ] and [z] are possible markers (cf. Labov 2001) of Barcelonan CCS speech. Having interviewed 12 speakers of CCS and monolingual (Madrid) Spanish ages 27–41 regarding their awareness of linguistic features of Catalanized Spanish, the only phonetic feature named by all speakers was [ɬ], suggestive of a stronger status as a linguistic stereotype (cf. Labov 2001) than intervocalic [z], which was named only by CCS speakers. Madrid speakers commented that Catalanized Spanish, and in particular [ɬ], sounded ugly, uneducated, and rough. Moreover, three CCS speakers admitted feeling an "obligation to correct or adjust their [Spanish] pronunciation when talking in public…" (Sinner 2002, 163, 165–166).

Wesch (1997) examined possible effects of age and social class on the frequency of [ɬ] and [z] in Barcelonan CCS. Having recorded spontaneous speech samples from 24 CCS speakers, Wesch (1997, 296, 298) reported that in the speech of Barcelonans ages 25–55, [ɬ] was least frequent in the speech of the younger speakers. These data can be interpreted as evidence of a possible linguistic change in progress whereby CCS [ɬ] is becoming less frequent over time, perhaps following as an effect of a social stigma associated with [ɬ] (cf. Sinner 2002). In contrast to [ɬ], [z] was reported as a frequent variant, though unlinked to any social factors (Wesch 1997, 296).

To synthesize, both [ɬ] and [z] are reported as present in the spontaneous speech of Barcelonan CCS speakers, albeit to different degrees. It seems the case that [ɬ] is a more salient feature of CCS than [z], carrying negative social value and

encouraging an innovative pattern of avoidance by Barcelonan youths. Still, the aforementioned studies merely describe [ɬ] and [z] as either generally present or absent, without empirical acoustic analysis. Also, linguistic factors that condition [ɬ] and [z] usage have not received attention. This study accordingly offers a quantitative and empirical examination of the extent to which [ɬ] and [z] have entered formal registers of Barcelonan CCS, and explores linguistic factors that condition each variant's usage as mediated by Catalan usage and exposure. We address these research questions:

1. How present are [ɬ] and [z] in formal registers of Barcelonan CCS speech?
2. What linguistic factors condition their usage?
3. To what extent are these linguistic factors as well as the overall usages of [ɬ] and [z] mediated by Catalan dominance?

3. Methodology

3.1 Subject population

The data analyzed for this study came from a select subset of the Barcelonan CCS-speaking population, namely twenty 18–27 year old female speakers of middle social class background. The decision to examine the speech of this specific population was grounded in standard sociolinguistic principles of language variation and change (cf. Labov 2001). To begin, it is widely accepted that "…women are the principal innovators in the process of [linguistic] change… [in that] women conform more closely than men to sociolinguistic norms that are overtly prescribed but conform less than men when they are not" (Labov 2001; 293–294; additionally Chambers 2004, 352). Since non-standard linguistic variants are typically avoided in formal speech registers (Tagliamonte 2012, 34), we can examine women's speech as a means of gathering maximally conservative estimates of the degree of [ɬ] and [z] usage in modern Barcelonan CCS.

The speech patterns of youth speakers is useful for making inferences about the most current language trends, since in the typical comparison of youth speakers' speech to that of older generations (cf. Bailey, Wikle, Tillery & Sand 1991; Chambers 2004), younger speakers' speech has been shown to reflect the most contemporary trajectories of language variation or a given change in progress (Tagliamonte 2012, 43–45). Furthermore, the focus on middle class speech also reflects Labovian principles of language variation and change, namely in that in cases of linguistic change from above, the middle social class often shows instances of exaggerated linguistic innovation extending beyond the norms of the higher social class, while in cases of linguistic changes from below, change is believed to

originate within a central social group within the interior of the socioeconomic hierarchy (Labov 2001, 188, 275).

In summary, following the aforementioned results of Sinner (2002) concerning differences in overt social value between Barcelonan CCS [ɬ] and [z] (indicative of the presence of an overtly prescribed norm for /l/ production but not for /s/), usage patterns of [ɬ] and [z] are expected to differ, particularly in the speech of young Barcelonan women of the middle class, and we have tailored our subject population to reflect this dynamic group of language users. Having restricted our scope to the speech of this particular group of speakers, however, usage patterns of [ɬ] and [z] will be examined according to a social construct of language dominance, which is detailed in Section 3.3.1.

3.2 Instruments and data collection

The first instrument in this study was a sociodemographic questionnaire (adapted from Pieras [1999]) regarding participants' family background and personal information, education, occupation, language use, and self-reported competence in Spanish and Catalan.

The second instrument was a Spanish recorded reading. Participants were asked to carefully read aloud, in their best Spanish pronunciation, a series of 162 two-word phrases that featured numerous linguistic contexts of /l/ and /s/, controlled in their distribution with respect to the linguistic factors detailed in Section 3.3.2 and placed within a carrier phrase translating to "I say ___ for you." To avoid confounds of vowel height on /s/ production (cf. File-Muriel & Brown 2011), /s/ tokens were surrounded by two low /a/ vowels, reported to least affect the energy frequencies of fricative segments (Quilis 1981, 235). Although data elicited from a word-reading task are not directly comparable with spontaneous speech, they reveal valid insights as to the kinds of variants produced in a formal speech style, and will undershoot actual production rates in natural speech.[2]

Participants were recruited by means of flyers posted at Barcelona university campuses. Participants were individually recorded in separate experimental sessions using an SE50 Samson head-mounted condenser microphone and an H4n Zoom digital recorder in an audiometric booth in the phonetics laboratory at the *Universitat Autònoma de Barcelona* or in an empty classroom at the *Universitat de Barcelona* or *Universitat Pompeu Fabra*.

2. A brief third instrument, a Catalan recorded reading of 12 short phrases, was administered at the end of the experimental session to confirm that all participants systematically produced Catalan [ɬ] and [z], permitting the analysis of CCS [ɬ] and [z] production as subject to transfer from Catalan.

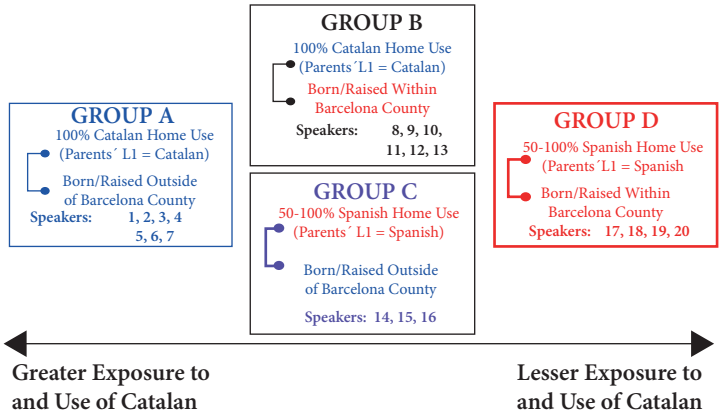

Figure 1. Hierarchical social construct of catalan exposure and use as crude language dominance measure

3.3 Independent variables

3.3.1 *Social factor groups*

To assess effects of language dominance on CCS /l/ and /s/ production, participants were classified into four groups based on a gradient, crude measure of language dominance, namely a social construct of degree of exposure to and use of Catalan (Figure 1). Home language use (while growing up) was selected for incorporation into the social construct because it is intrinsically linked to Catalan usage. In the absence of a formalized proficiency measure, home language use serves as an indicator of language dominance (for example, age of acquisition of Catalan as either 0–3 years old [L1] or 6+ years old [L2, acquired in school]). Catalan home usage was divided into two extremes: 100% Catalan use vs. 50–100% Spanish use.[3] Additionally, locality of residence was selected for incorporation into the social construct because it is intrinsically linked to the amount of Catalan that subjects are exposed to from their community. Barcelona County (*Comarca Barcelonès*) is renowned as the area of Catalonia with the highest presence of Spanish, attributed to massive waves of monolingual Spanish speakers from Spain and Latin America during the mid-20th century (Lleó, Cortés & Benet 2008, 186). Census polls (cf. IEC 2010) indicate dramatically higher daily usage rates of Catalan in smaller

3. Though it would have been ideal to juxtapose a group of 100% Catalan and 100% Spanish home users, we could not find enough subjects that used exclusively Spanish at home, so the latter group includes individuals with at least 50% Spanish home use. Note that parent-L1 is controlled within each group: the 100% Catalan home use group has two Catalan-L1 parents, and the 50–100% Spanish home use group has two Spanish-L1 parents (that were born and raised in monolingual regions of Spain and immigrated to Catalonia as adults).

counties and towns just outside of Barcelona County than in the urban capital (consistent with self-reported language use estimates by each participant in the sociodemographic questionnaire), and accordingly we divided speakers into two groups of locality of residence: within Barcelona County and outside of Barcelona County.

3.3.2 *Linguistic factor groups*

The effects of five linguistic factors on /l/ and /s/ production were explored: (1) syllable position, (2) stress, (3) accentual unit, (4) morpheme class, and (5) cognate status. Two additional linguistic factors pertained solely to /l/ production: (6) following consonant place of articulation, and (7) prior vowel frontedness.

Syllable position, with respect to /l/ production, was coded as either word-medial (e.g. *huelga* 'strike') or word-final (e.g. *papel grande* 'big paper'). CCS [ɬ] production in Pieras (1999), Serrano Vázquez (1996), Vann (2001), and Wesch (1997) was reported exclusively in word-final coda tokens, and as such it is unclear if velarization is equally present in word-medial coda tokens. Both positions nonetheless featured coda /l/ as opposed to a syllable onset /l/, reflecting cross-linguistic (e.g. American English, Chicano English, Catalan) findings of weaker lateral velarization for /l/s in syllable onset position than coda position (cf. Hofwegen 2009; Recasens 1986, 1993; Sproat & Fujimura 1993). With respect to /s/ production, syllable position (intervocalic in all cases) was coded as either word-medial (e.g. *payasa contenta* 'happy clown') or word-final (e.g. *caminarás aquí* '2.s. will walk here'). CCS [z] production in Pieras (1999), Serrano Vázquez (1996), Vann (2001), and Wesch (1997) was reported exclusively in word-final tokens, and as such it is unclear if voicing is equally present in the word-medial position. We may expect that CCS voicing be more strongly disfavored word-medially because this position is a site of phonemic voicing contrast in Catalan, whereas the word-final position, being a site of purely phonetic voicing for both languages (albeit systematic in Catalan), may constitute a better candidate for cross-linguistic phonetic transfer from Catalan to CCS.

Stress, with respect to /l/ production, was coded as either stressed (e.g. *algo* /álgo/ 'something') or unstressed (e.g. *algodón* /algodón/ 'cotton') based on the vocalic nucleus of the syllable containing /l/. With respect to /s/ production, since sequences of vowel + /s/ + vowel contain two vowels permitting four possible combinations of stress (e.g. (/ásá/, /ása/, /asá/, and /asa/), stress was considered as two factor groups: preceding vowel stress (e.g. *serás apto* [se.ɾá.sáp.to] '2.s. will be suitable' – stressed; *fueras apto* [fwé.ɾa.sáp.to] 'that 2.s. were suitable' – unstressed) and following vowel stress (e.g. *fueras* apto [fwé.ɾa.sáp.to] 'that 2.s. were suitable' – stressed; fueras animado [fwé.ɾa.sa.ni.má.ðo] 'that 2.s. were animated' – unstressed). Word-medial /s/ tokens exhibited only /ása/ and /asá/ stress patterns, rendering the two factor groups redundant, and were coded only for following vowel stress. Hualde (2005, 244) notes that stressed syllables have longer durations

in Spanish than unstressed syllables, which, regarding /l/ production, would facilitate a more independent series of articulatory gestures for the vowel preceding /l/, /l/, and the following consonant (cf. Browman & Goldstein 1991). Thus, for stressed-syllable tokens like *algo*, /l/ production was expected to be less susceptible to coarticulation effects than for unstressed-syllable tokens such as *algodón*. Similarly, for /s/ production, we predicted that voicing would be most disfavored across stressed vowel + /s/ + vowel sequences, as these sequences most facilitate the independent series of articulatory gestures necessary for vocal fold adduction for each voiced vowel and vocal fold abduction for voiceless [s]. That is, a sequence such as /ásá/ would maximally disfavor voicing since the vowels adjacent to /s/ are maximally long (compared to their unstressed counterparts), facilitating vocal fold abduction for /s/ independently from each vowel's vocal fold adduction.

Accentual unit, with respect to /l/ production, was coded as either singular for /l/ tokens in a function word + noun phrase (e.g. *aque**l** caldo* 'that soup') or double for /l/ tokens in a noun + adjective phrase (e.g. *hote**l** caro* 'expensive hotel'). With respect to /s/ production, accentual unit was coded as either singular for /s/ tokens in a function word + noun phrase (e.g. *la**s** aguas* 'the waters') or double for /s/ tokens in a noun/verb + adjective/adverb phrase (e.g. *aguas ácidas* 'acidic waters'). The function words used as experimental stimuli (e.g. determiners, demonstrative pronouns, prepositions) lack lexical stress, and thus form a single accentual unit (or stress domain) with the following lexically stressed noun/adjective/adverb, which following Browman & Goldstein (1991) and Hualde (2005) would facilitate greater coarticulation of /s/ and /l/ with adjacent segments (Hualde 2005, 233–234, 244). Note that this factor group applied only to word-final /l/ and /s/ tokens, and was additionally motivated by Campos-Astorkiza (forthcoming), who hypothesized for Spanish that single accentual units would likely show differences in /s/ voicing assimilation from double accentual units. This hypothesis stemmed from the unexpectedly similar voicing assimilation rates observed between /s/ in single accentual units (across a word boundary) and word-internal /s/ (word boundary absent) in her Peninsular Spanish data.

Morpheme class, with respect to /l/ production, was coded as either derivational suffix (e.g. *plante**l** grande* 'big plant farm') or lexical stem (e.g. *hote**l** grande* 'big hotel') based on if /l/ was part of a suffix or not. With respect to /s/ production, morpheme class was coded as verbal inflection (e.g. *fuera**s** ágil* 'that 2.s. were agile'), nominal inflection (e.g. *amiga**s** altas* 'tall friends'), or lexical stem (e.g. *atla**s** amplio* 'wide atlas'). Note that this factor group applied only to word-final /l/ and /s/ tokens, and was motivated by general principles of sound change, in that "lexical representations, including phonological representations, are linked in a connectionist network to other lexical representations, allowing generalizations to emerge which form the basis for the spread of the sound change to new words depending on phonetic or morphological resemblance" (Phillips 2006, 182).

Cognate status, with respect to /l/ production, was coded as either lesser cognate (e.g. *papel pálido* 'pale paper' – Catalan *paper* [pəpé]) or greater cognate (*hotel pálido* '' – Catalan *hotel* [utɛ́ɬ]) based on if the word featured /l/ in both languages. Note that following Brown & Harper (2009) and Costa, Santesteban & Caño (2005), greater cognate tokens were expected to most facilitate CCS [ɬ] production in the sense that these tokens, within an interconnected bilingual lexicon, trigger a parallel lexical activation of a Catalan cognate containing the velarized lateral category (compared to lesser cognate items that in Catalan lack /l/ entirely). With respect to /s/ production, cognate status was coded as either lesser cognate (e.g. *casa grande* 'big house' – Catalan *casa* /káza/) for words which in Catalan did not feature /s/ or greater cognate (e.g. *masa grande* 'big dough' – Catalan *massa* /mása/) for words that in both languages featured /s/. Similarly, lesser cognate tokens were expected to facilitate CCS [z] production, since these are precisely the tokens that feature /z/ as a source for direct transfer to CCS via parallel lexical activation.

The final two linguistic factors, exclusive to /l/ production, were following consonant place of articulation and prior vowel frontedness. The former was coded as bilabial (e.g. *aquel baño* 'that bathroom'), dental (e.g. *aquel dato* 'that detail'), or velar (e.g. *aquel gato* 'that cat'), whereas the latter was coded as front (e.g. belga *'Belgian female'*) or non-front (e.g. *talco* 'talc'). Both coding schemes were motivated by traditional Catalan literature on Catalan /l/ coarticulation that shows strongest velarization degrees in contexts of an adjacent non-front vowel and either a velar or labial consonant (cf. Recasens 1986, 102; 1991, 307; 1993, 178–179; Wheeler 1979, 306–308).

3.4 Analysis of dependent variables

Audio data consisting of 3,240 tokens of /l/ and /s/ were submitted to acoustic analysis as *wav* files using Praat. Lateral velarization was measured as a function of second formant frequency (F2), taken from the midpoint of each lateral following manual delimitation of the /l/ segments' boundaries (cf. Simonet 2010). Fricative voicing was measured as a function of the percentage of an /s/ segment's duration that was voiced (cf. File-Muriel & Brown 2011; Gradoville 2011). Voicing durations were measured as portions of the /s/ segment with a fundamental frequency, a voice bar at the bottom of the spectrogram, and glottal pulses (cf. Schmidt & Willis 2011). There were 544 /l/ and /s/ productions with speaker disfluencies or uninterpretable spectrographic images that were discarded, the majority of which were word-final /s/ tokens (e.g. *aguas ácidas* 'acidic waters') in which a pause between the words disrupted the desired /asa/ sequence.

The coding scheme for [ɬ] vs. [l] and [z] vs. [s] took into account actual distributions of F2 values (for /l/) and voicing percentages (for /s/), as well as velarization and voicing thresholds reported in past empirical research in Catalan and Spanish

(and was coherent with the author's auditory judgment). For /l/, Pieras (1999, 220), Prieto (2004, 204) Quilis (1981, 276), Recasens (1986, 95, 102; 1991, 306; 2004, 599), and Recasens et al. (1995, 42) report that [ɬ] and [l] tend to exhibit F2 frequencies of below and above 1500hz, respectively. Our /l/ production data, however, was not distributed in a strictly bimodal fashion across the 1500hz threshold, and instead tended to aggregate as a group of productions with F2 frequencies below 1200hz and another above 1500hz. We accordingly coded as [ɬ] /l/ tokens with F2 values of 1200hz or below, and as [l] /l/ tokens with F2 values of 1500hz or above.

Regarding /s/ coding, there are considerably fewer empirical studies that report voicing thresholds between [s] and [z] for Spanish and/or Catalan. File-Muriel & Brown (2011, 224–225) note that the majority of empirical research on /s/ lenition phenomena in Spanish establish thresholds between /s/ variants impressionistically, using native speaker judges to code /s/ productions by ear. For Catalan, this coding method was employed most recently by Benet, Cortés & Lleó (2012, 396–397) to distinguish between [z] and [s] productions. Still, we have found one example of a voicing threshold for Spanish in Schmidt & Willis (2011, 6), who reported that [z] productions in Spanish tended to exhibit voicing durations lasting 60% or more of the segments' durations. Our /s/ production data, however, was not distributed in a strictly bimodal fashion across the 60% voicing threshold, and instead tended to aggregate as a group of productions with voicing percentages below 30% and another above 60%. We accordingly coded as [z] /s/ tokens with voicing durations of 60% or above, and as [s] /s/ tokens with voicing durations of 30% or below.

Figures 2 and 3 offer illustrative comparisons between /l/ and /s/ productions coded as [ɬ]/[z] and [l]/[s], from speakers 14 and 1. For ease of comparison, /l/ and /s/ productions have been spliced adjacent to one another, separated from their original phrasal contexts.

The remaining total of 2,491 tokens coded as [s], [z], [l], and [ɬ] were submitted to variable-rule logistic regressions using Goldvarb. Aside from being able to handle particularly unbalanced datasets (typical of sociolinguistic speech data), Goldvarb also offers the advantage of being able to handle nested independent variables, that is, independent variables that apply to only a subset of the data (i.e. morpheme class applied only to word-final /l/ and /s/ tokens, etc.) (Roy 2013, 265; Tagliamonte 2012, 156).

4. Results

4.1 Production of [ɬ] and [z] by individual speaker

A descriptive statistical analysis of each participant's frequency of [ɬ] and [z] production is illustrated in Figure 4, organized by social construct group.

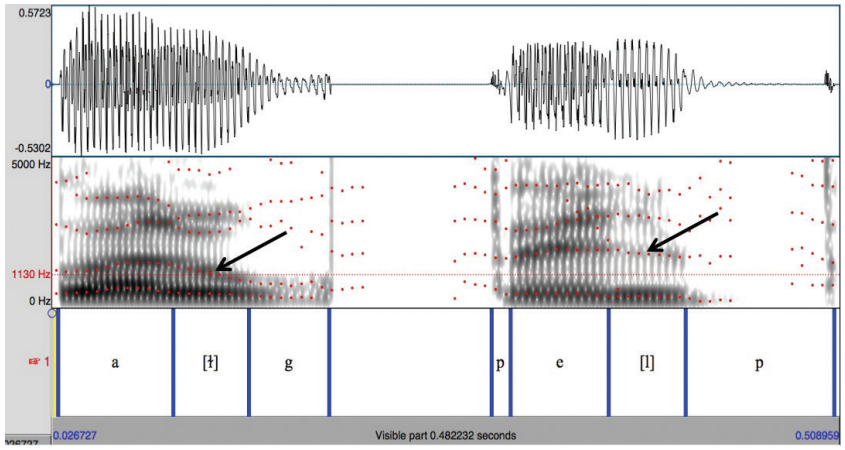

Figure 2. Speaker 14 renditions of *alguacil* 'sheriff' (1170hz) and *papel pálido* 'pale paper' (1800hz)

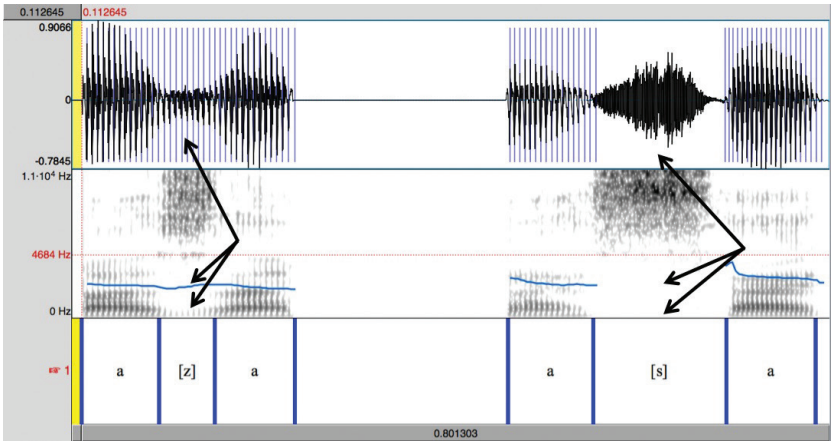

Figure 3. Speaker 1 renditions of *las amigas* 'the friends' (100% voiced) and *casaca* 'jacket' (3% voiced)

Note that while speakers from group A overall showed the most consistent and highest usage of both [ɬ] and [z], all groups show within-group heterogeneity, which is to be expected particularly for non-standard variants. At least one speaker from each group (e.g. speakers 7, 9, 16, and 17) produced [ɬ] and [z] as little as 1% or less of the time, while other speakers (e.g. speakers 1, 5, 6, and 19) produced either [ɬ] or [z] more than 75% of the time. It should be highlighted that whereas [ɬ] production rates above 33% were exhibited only by speakers in group A, [z] production rates above 33% were additionally produced by select members of group D. Differences between speakers 19 and 20, who

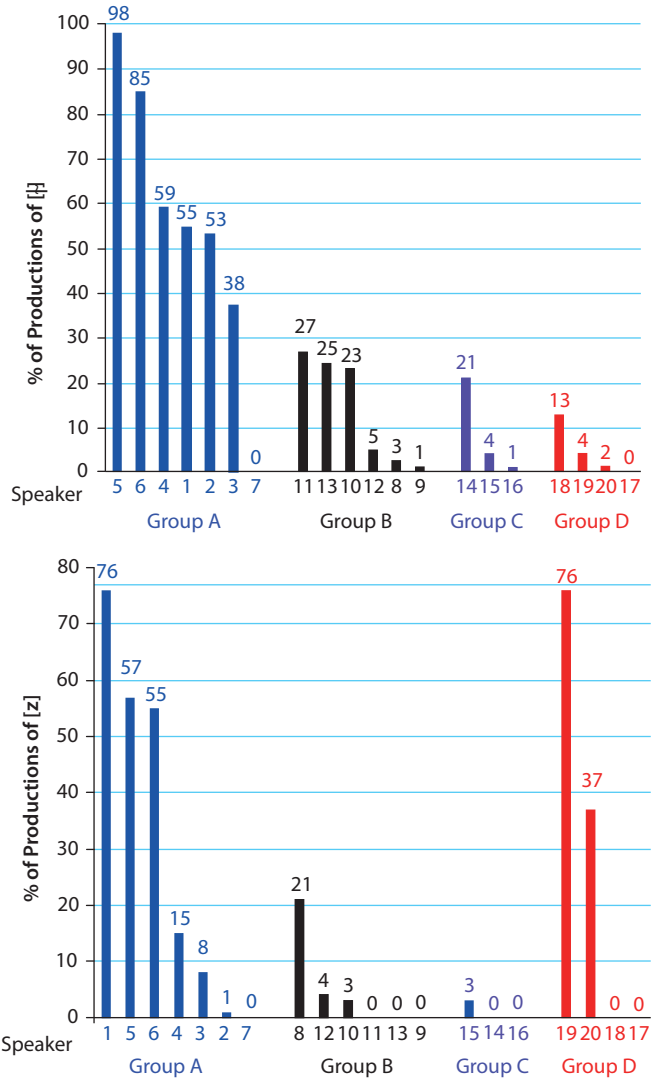

Figure 4. Individual speakers' frequency of [ɬ] and [z] production by catalan exposure and use group

rather frequently produced [z] in their careful Spanish readings, and speakers 17 and 18, who categorically favored [s] over [z], are not easily attributable to differences in Catalan proficiency. Self-reported Catalan proficiency (collected in the sociodemographic questionnaire) did not largely vary between the [z]- and [s]-users of group D, and recall that all participants did produce [z] when reading in Catalan.

4.2 Social and linguistic constraints on [ɬ] and [z] production

Results from Goldvarb analyses are reported in Tables 1, 2, 3, and 4 in the form of weighted, hierarchical rankings of social and linguistic factors that favor the production of [ɬ] and [z] over [l] and [s], respectively. Social and linguistic factor groups were separated in statistical iterations, following Tagliamonte (2012, 129). Factor weights of above 0.5, in bold, indicate that [ɬ] or [z] is favored by an independent variable level (e.g. group A speech, unstressed syllable context, etc.), and range values indicate the relative strength of effect of each independent variable. As Goldvarb cannot run inferential statistics without some degree of variation in dependent variable outcomes for each independent variable, categorical data were submitted to a Fisher's Exact Test in SPSS, which is a conservative version of the Chi-Square test more appropriate for categorical data (Gorman & Johnson 2013, 219–220). These results have been integrated into each table as p-value statistics instead of factor weights, and models for which only Fisher's Exact Tests reached significance do not generate model significance p-values or log likelihoods.

Table 1. Varbrul logistic regression weights for [ɬ] production by social construct group

Total N: 1129		**Corrected Mean: .154**	
Social construct group	**Weight**	**% [ɬ]**	**N**
A	**.83**	47.3	319
B	.43	12.1	371
C	.34	8.5	213
D	.24	5.3	226
RANGE	*59*		

Log Likelihood = –466.340, Logistic Regression Model Significance: $p < .001$

Table 2. Varbrul logistic regression weights for [z] production by social construct group

Total N: 1362		**Corrected mean: .121**	
Social construct group	**Weight**	**% [z]**	**N**
A	**.77**	32.1	458
B-C †	.21	3.5	634
D	**.74**	28.5	270
RANGE	*56*		

Log Likelihood = –544.395, Logistic Regression Model Significance: $p < .001$

Table 3. Logistic regression weights for [ɬ] production by linguistic factor per social construct group

	Group A		
Total N: 319			**Corrected Mean: .478**
Prior Vowel Frontedness	**Weight/Fisher's Exact Test p-value**	**% [ɬ]**	**N**
Non-Front	**.78**	79.1	86
Front	.38	35.6	233
RANGE	*40*		
Following consonant place of articulation			
Velar	**.54**	54.6	238
Dental	.41	26.8	41
Bilabial	.39	25.0	40
RANGE	*15*		
Cognate Status			
Greater Cognate (Catalan /l/)	**.55**	52.8	163
Lesser Cognate (No Catalan /l/)	.45	41.7	156
RANGE	*10*		

Log Likelihood = −193.544, Logistic Regression Model Significance: p = .035
Not selected as significant: Stress, Accentual Unit, Morpheme Class, Syllable Position

	Group B		
Total N: 371			**Corrected Mean: .051**
Following consonant place of articulation			
Velar	**p < .001**	19.2	234
Non-Velar		0.0	137
Prior Vowel Frontedness			
Non-Front	**.98 ††**	67.9	53
Front	.35	2.8	318
RANGE	*63*		

Log Likelihood = −74.209, Logistic Regression Model Significance: p < .001
Not selected as significant: Stress, Accentual Unit, Morpheme Class, Syllable Position, Cognate Status

	Group C		
Total N: 213			**Corrected Mean: .028**
Following consonant place of articulation			
Velar	**p < .001**	12.8	141
Non-Velar		0.0	72

(*Continued*)

Table 3. (Continued)

	Weight/Fisher's Exact Test p-value	% [ɬ]	N
Prior Vowel Frontedness			
Non-Front	**.98** ††	53.6	28
Front	.36	1.6	185
RANGE	*62*		

Log Likelihood = −34.677, Logistic Regression Model Significance: p < .001
Not selected as significant: Stress, Accentual Unit, Morpheme Class, Syllable Position, Cognate Status

Group D

Total N: 226			**Corrected Mean: .015**
Following consonant place of articulation			
Velar	**p = .002**	9.0	133
Non-Velar		0.0	93
Prior Vowel Frontedness			
Non-Front	**p < .001**	32.4	37
Front		0.0	189

Not selected as significant: Stress, Accentual Unit, Morpheme Class, Syllable Position, Cognate Status

Table 4. Logistic regression weights for [z] production by linguistic factor per social construct group

Group A

Total N: 458			**Corrected Mean: .405**
Syllable Position	**Weight/Fisher's Exact Test p-value**	**% [z]**	**N**
Word-Final Intervocalic	**p < .001**	42.1	349
Word-Medial Intervocalic		0.0	109
Accentual Unit			
Double	**.52**	43.5	271
Single	.43	37.2	78
RANGE	*9*		
Following Vowel Stress			
Unstressed	**.53**	34.2	230
Stressed	.48	30.0	228
RANGE	*5*		

(*Continued*)

Table 4. (Continued)

	Weight/Fisher's Exact Test p-value	% [z]	N
Preceding Vowel Stress			
Unstressed	**.52**	42.7	213
Stressed	.47	41.2	136
RANGE	*5*		

Log Likelihood = -256.265, Logistic Regression Model Significance: p < .001 Not selected as significant: Morpheme Class, Cognate Status

Groups B–C †

Total N: 634		**Corrected Mean: .029**	
Syllable Position			
Word-Final Intervocalic	**.60**	4.3	491
Word-Medial Intervocalic	.19	0.7	143
RANGE	*41*		

Log Likelihood = -92.694, Logistic Regression Model Significance: p=.017 Not selected as significant: Preceding Vowel Stress, Following Vowel Stress, Accentual Unit, Morpheme Class, Cognate Status

Group D

Total N: 270		**Corrected Mean: .361**	
Syllable Position			
Word-Final Intervocalic	**.69**	36.4	209
Word-Medial Intervocalic	.06	1.6	61
RANGE	*63*		
Accentual Unit			
Double	**.53**	39.0	164
Single	.39	26.7	45
RANGE	*14*		

Log Likelihood = -157.187, Logistic Regression Model Significance: p = .006 Not selected as significant: Preceding Vowel Stress, Following Vowel Stress, Morpheme Class, Cognate Status

† Groups B and C are reported together for [z] because their separation as 2 groups did not significantly contribute to the model's ability to account for variance in [z] production, as determined by a Chi-Square test (p > .05) evaluating log likelihood differences between the model with groups B and C separated and the model with groups B and C combined (cf. Young & Bayley 1996).
†† Inflated effects (weights of .98) of prior vowel frontedness for groups B and C can be accounted for by technical reasons. Since Fisher's Exact Tests were needed to analyze the factor of following consonant place of articulation, the regression analyses were forced to account for variance in [ɬ] production solely with prior vowel frontedness, resulting in its inflated factor weight.

Summarizing the results presented in Tables 1 and 2, Catalan dominance accounts for a significant portion of the variance in [ɬ] and [z] production, in that [ɬ] is favored by group A speakers and progressively disfavored from groups B to D, while [z] is favored by both groups A and D. Summarizing the results presented in Tables 3 and 4, regarding [ɬ] production, all groups show a significant sensitivity to coarticulation effects of a prior vowel and following consonant, whereas only group A shows an additional sensitivity to cognate status. Regarding [z] production, whereas all groups show a significant sensitivity to syllable position, groups A and C both show a significant sensitivity to accentual unit, and only group A shows an additional significant sensitivity to the stress of each of the two vowels in the /asa/ sequence.

5. Discussion

With respect to the presence of [ɬ] and [z] in formal registers of Barcelonan CCS, our results suggest that both variants are in fact present, at times as extremely competitive (nearly categorical) variants of /l/ and /s/. Their presence in the careful speech of youth Barcelonan females suggests that both remain vitalic variants of /l/ and /s/ in casual speech, and that neither has become so heavily stigmatized that it has been completely abandoned by younger CCS speakers, contrary to data collected by Wesch (1997) for [ɬ]. Still, neither variant is used preferentially over [l] or [s] (i.e. usages of 50% or more in Tables 1 and 2) on average by any speaker group, suggesting that neither variant has, in the last 16 years since Wesch (1997), become a normative, majority variant (fully adopted into formal speech registers), particularly in Barcelona County (groups B and D). Regarding smaller villages outside Barcelona County, Figure 4 shows that five out of the seven group A speakers use [ɬ] more often (for some, considerably more often) than [l]. This may constitute a preliminary validation of Sinner's (2002) comments regarding an association of [ɬ] usage with rurality, but when considered alongside the rather dissimilar [ɬ] usage between groups A and C, it indicates that lateral velarization is quite strongly linked to a high degree of exposure to and use of Catalan (also evidenced by the downward sloping [ɬ] production from groups A to D). The same cannot be claimed as strongly for [z], which although showed highest usage rates by group A speakers, nonetheless also appeared as a relatively competitive /s/ variant for select speakers in group D with the least exposure to and use of Catalan.

With respect to linguistic factors conditioning [ɬ] production, all four speaker groups showed sensitivity to coarticulation effects in the same direction, favoring velarization in contexts of an adjacent non-front vowel and velar consonant

(as predicted by models of gestural phonology, cf. Browman & Goldstein 1991). However, cognate status only conditioned [ɬ] production for one of the four speaker groups. For speakers of group A, with the highest exposure to and usage of Catalan, cognate words featuring /l/ in both languages favored lateral velarization, whereas speakers with lesser exposure to and usage of Catalan were not sensitive to cognate effects. This observed cognate effect, precisely mediated by our crude measure of Catalan dominance (as predicted within an interconnected bilingual lexicon, cf. Brown & Harper [2009]; Costa et al. [2005]), lends support to the classification of [ɬ] as a contact variant, originally entering Spanish via phonetic transfer from Catalan in speakers for whom Catalan is a sufficiently dominant language. Regarding the strength of linguistic factor constraints (or range hierarchies in Table 3) across social construct groups for [ɬ] production, group A shows the strongest sensitivity to prior vowel frontedness, followed by following consonant place of articulation and lastly cognate status, whereas the other groups show comparable sensitivities to each type of coarticulation effect.

With respect to linguistic factors conditioning [z] production, all four speaker groups showed sensitivity to syllable position in the same (predicted) direction, nearly categorically favoring [s] over [z] word-medially. This particularly strong constraint barring voicing word-medially lends support to [z] production in CCS as a contact phenomenon, since this context is the site of phonemic voicing contrast in Catalan.[4] To this end, the unexpected finding of a lack of effect of cognate status on [z] production is likely due to the strength of the syllable position constraint. Cognate status was only applicable to word-medial tokens, only 0.6% (3 tokens) of which were produced as [z], although all 3 were (as predicted) lesser cognate tokens. Additionally, while groups A and D (the only groups with speakers that somewhat consistently produced [z]) showed sensitivity to accentual unit, only group A showed sensitivity to effects of stress, favoring [z] production across sequences of unstressed /asa/, as predicted by models of gestural phonology (cf. Browman & Goldstein 1991).[5] As was the case for [ɬ] production, speakers that produce [z] most consistently show the most complex patterns of linguistic

4. Moreover, it is critical to note that voicing differences between word-final and word-medial intervocalic positions in the casual speech of monolingual Spanish varieties (cf. Schmidt & Willis 2011; Torreira & Ernestus 2012) do not come close to approaching the nearly categorical constraint attested in the present study with Catalan-Spanish bilinguals.

5. Groups A and D unexpectedly favored [z] in contexts of a double accentual unit. Since examples of prototypical CCS [z] production cited by Pieras (1999, 212), Serrano Vázquez (1996, 379), Vann (2001, 124), and Wesch (1997, 296) are all single accentual unit /s/ tokens, it is possible that this context (especially determiner + noun, as in *los amigos* 'the friends') is the most salient context for [z] production, and thus appears least in formal speech registers in

constraints. Range hierarchies in Table 4 show that syllable position is the strongest constraint across all groups, followed by accentual unit for groups A and D, and lastly stress effects for group A.

The intriguing divergence in group D's usage patterns between [ɬ] and [z] production (not in raw production rates, but rather the patterning with group A speakers) warrants further investigation.[6] If [ɬ] is a considerably salient marker of CCS speech, reaching the status of linguistic stereotype (cf. Sinner 2002), then it may be the case that less salient CCS features, such as [z], would not be as restricted in CCS speech and could have greater potential to be adopted by speakers of lesser Catalan dominance. A larger-scale apparent time study, which, in addition to incorporating additional social factors such as gender, age, and social class, also collected perception and attitudinal data on [ɬ] and [z] in formal and casual Barcelonan speech, is needed (and currently underway) to fully account for the present data trends.

6. Conclusion

The present study has aimed to provide an in-depth, quantitative approach to examining [ɬ] and [z] production in Barcelonan CCS. We have provided evidence that both [ɬ] and [z] exist (even as competitive variants) in formal speech registers, which underestimates their actual presence in natural (spontaneous) CCS speech. Both variants are conditioned by a set of linguistic factors, most notably adjacent segment coarticulation effects for [ɬ] and syllable position effects for [z]. Additional linguistic constraints, mediated by Catalan dominance, were found to account for significant portions of variance in [ɬ] production in select speakers outside of Barcelona County, whereas additional linguistic constraints on [z] production were found more generally in the speech of speakers with the highest

which non-standard variants are typically most avoided. Further investigation incorporating the perception of voicing is certainly needed to confirm this speculation.

6. A direct comparison between average [ɬ] and [z] production in Figure 4 does not lend itself to suggest that one variant is more frequent than the other. Recall that all speaker groups showed a near categorical preference for [s] over [z] word-medially, whereas /l/ production was not constrained in such a broad and categorical manner by any of the linguistic factors studied. Given this asymmetric constraint pattern between the two dependent variables, the generally lower usage rates for [z] compared with [ɬ] in Figure 4 are unsurprising. Moreover, recall (from Section 3.4.) that a majority of excluded tokens (due to speaker disfluencies) were /s/ tokens in word-final position, precisely the position that permitted over 99% of observed [z] productions.

voicing rates. Future research would benefit from the incorporation of perception and attitudes tests associated with CCS variants like [ɬ] and [z] in order to better understand the social dynamics underlying the linguistic behaviors observed in this and prior CCS studies.

References

Bailey, Guy, Tom Wikle, Jan Tillery, and Lori Sand. 1991. "The Apparent Time Construct." *Language Variation and Change* 3 (3): 241–264. DOI: 10.1017/S0954394500000569

Benet, Ariadna, Susana Cortés, and Conxita Lleó. 2012. "Devoicing of Sibilants as a Segmental Cue to the Influence of Spanish onto Current Catalan Phonology." In *Multilingual Individuals and Multilingual Societies*, ed. by Kurt Braunmüller, and Christoph Gabriel, 391–404. Philadelphia, PA: John Benjamins. DOI: 10.1075/hsm.13.25ben

Bonet, Eulàlia, and Maria-Rosa Lloret. 1998. *Fonologia catalana* [Catalan Phonology]. Barcelona: Ariel.

Browman, Catherine, and Louis Goldstein. 1991. "Gestural Structures: Distinctiveness, Phonological Processes, and Historical Change." In *Modularity and the Motor Theory of Speech Perception: Followings of a Conference to Honor Alvin M. Liberman*, ed. by Ignatius M. Mattingly, and Michael Studdert-Kennedy, 313–338. Hillsdale, NJ: Lawrence Erlbaum.

Brown, Esther, and David Harper. 2009. "Phonological Evidence of Interlingual Exemplar Connections." *Studies in Hispanic and Lusophone Linguistics* 2 (2): 257–274.

Campos-Astorkiza, Rebeka. Forthcoming. 2011. "Sibilant Voicing Assimilation in Peninsular Spanish as Gestural Blending." In *Romance Linguistics*, ed. by Marie-Hélène, Éric Mathieu, and Shana Poplack, 26p. Philadelphia, PA: John Benjamins.

Chambers, Jack. 2004. "Patterns of Variation Including Change." In *The Handbook of Language Variation and Change*, ed. by Jack Chambers, Peter Trudgill, and Natalie Schilling-Estes, 349–372. Oxford: Blackwell.

Costa, Albert, Mikel Santesteban, and Agnès Caño. 2005. "On the Facilitatory Effects of Cognate Words in Bilingual Speech Production." *Brain and Language* 94: 94–103. DOI: 10.1016/j.bandl.2004.12.002

File-Muriel, Richard, and Earl Brown. 2011. "The Gradient Nature of s-Lenition in Caleño Spanish." *Language Variation and Change* 23: 223–243. DOI: 10.1017/S0954394511000056

Gorman, Kyle, and Daniel E. Johnson. 2013. "Quantitative Analysis." In *The Oxford Handbook of Sociolinguistics*, ed. by Robert Bayley, Richard Cameron, and Ceil Lucas, 214–240. Oxford: Oxford University Press.

Gradoville, Michael. 2011. "Validity in Measurements of Fricative Voicing: Evidence from Argentine Spanish." In *Selected Proceedings of the* 5th *Conference on Laboratory Approaches to Romance Phonology*, ed. by Scott M. Alvord, 59–74. Somerville, MA: Cascadilla.

Hofwegen, Janneke Van. 2009. "Cross-generational change in /l/ in Chicano English." *English World-Wide* 30 (3): 302–325. DOI: 10.1075/eww.30.3.04van

Hualde, José Ignacio. 2005. *The Sounds of Spanish*. Cambridge: Cambridge University Press.

Hualde, José Ignacio, Anxton Olarrea, Anna María Escobar, and Catherine Travis. 2010. *Introducción a la lingüística hispánica* [Introduction to Hispanic Linguistics]. 2nd edition. Cambridge: Cambridge University Press.

Institut d'Estadística de Catalunya (IEC). 2010. "Territori: Banc d'estadístiques de municipis i comarques [Territory: Bank of Statistics of Municipalities and Counties]." Idescat: Generalitat de Catalunya. ⟨http://www.idescat.cat⟩.

Labov, William. 2001. *Principles of Linguistic Change: Social Factors, 2*. London: Blackwell.

Lleó, Conxita, Susana Cortés, and Ariadna Benet. 2008. "Contact-Induced Phonological Changes in the Catalan Spoken in Barcelona." In *Language Contact and Contact Languages*, ed. by Peter Siemund, and Noemi Kintana, 185–212. Philadelphia, PA: John Benjamins.

Phillips, Betty S. 2006. *Word Frequency and Lexical Diffusion*. New York, NY: Palgrave Macmillan. DOI: 10.1057/9780230286610

Pieras, Felipe. 1999. Social Dynamics of Language Contact in Palma de Mallorca: Attitude and Phonological Transfer. Ph.D. Dissertation, Pennsylvania State University.

Prieto, Pilar. 2004. *Fonètica i fonologia: Els sons del català* [Phonetics and Phonology: The Sounds of Catalan]. Barcelona: Editorial UOC.

Quilis, Antonio. 1981. *Fonética acústica de la lengua española* [Acoustic Phonetics of the Spanish Language]. Madrid: Gredos.

Recasens, Daniel. 1986. *Estudis de fonètica experimental del català oriental central* [Experimental Phonetic Studies on East-Central Catalan]. Barcelona: Publicacions de L'Abadia de Montserrat.

Recasens, Daniel. 1991. *Fonètica descriptiva del català* [Descriptive Phonetics of Catalan]. Barcelona: Institut d'Estudis Catalans.

Recasens, Daniel. 1993. *Fonètica i fonologia* [Phonetics and Phonology]. Barcelona: Enciclopèdia Catalana.

Recasens, Daniel. 2004. "Darkness in [l] as a Scalar Phonetic Property: Implications for Phonology and Articulatory Control." *Clinical Linguistics and Phonetics* 18 (6–8): 593–603. DOI: 10.1080/02699200410001703556

Recasens, Daniel, Jordi Fontdevila, and Maria Dolors Pallarès. 1995. "Velarization Degree and Coarticulatory Resistance for /l/ in Catalan and German." *Journal of Phonetics* 23: 37–52. DOI: 10.1016/S0095-4470(95)80031-X

Roy, Joseph. 2013. "Sociolinguistic Statistics: The Intersection between Statistical Models, Empirical Data and Sociolinguistic Theory." In *Proceedings of Methods XIV: Papers from the Fourteenth International Conference on Methods in Dialectology*, ed. by Alena Barysevich, Alexandra D'Arcy, and David Heap, 261–275. Frankfurt: Peter Lang.

Schmidt, Lauren, and Erik Willis. 2011. "Systematic Investigation of Voicing Assimilation of Spanish /s/ in Mexico City." In *Selected Proceedings of the 5th Conference on Laboratory Approaches to Romance Phonology*, ed. by Scott M. Alvord, 1–20. Somerville, MA: Cascadilla. DOI: 10.1016/j.wocn.2010.10.002

Serrano Vázquez, María del Carmen. 1996. "Rasgos fonéticos del carácter interferencial en el castellano de una comunidad bilingüe" [Interferential phonetic features of the Spanish of a bilingual community]." *Estudios de Lingüística: Universidad de Alicante* 11: 365–383.

Simonet, Miquel. 2010. "Dark and Clear Laterals in Catalan and Spanish: Interaction of Phonetic Categories in Early Bilinguals." *Journal of Phonetics* 38: 663–678.

Sinner, Carsten. 2002. "The Construction of Identity and Group Boundaries in Catalan Spanish." In *Us and Others: Social Identities across Languages, Discourses and Cultures*, ed. by Anna Duszak, 159–186. Philadelphia, PA: John Benjamins. DOI: 10.1075/pbns.98.10sin

Sproat, Richard, and Osamu Fujimura. 1993. "Allophonic Variation in English /l/ and its Implications for Phonetic Implementation." *Journal of Phonetics* 21: 291–311.

Tagliamonte, Sali. 2012. *Variationist Sociolinguistics: Change, Observation, Interpretation*. Oxford: Wiley-Blackwell.

Torreira, Francisco, and Mirjam Ernestus. 2012. "Weakening of Intervocalic /s/ in the Nijmegen Corpus of Casual Spanish." *Phonetica* 69: 124–148. DOI: 10.1159/000343635

Vann, Robert E. 2001. "El castellà catalanitzat a Barcelona: Perspectives lingüístiques i culturals [The Catalanized Spanish of Barcelona: Linguistic and cultural perspectives]." *Catalan Review: International Journal of Catalan Culture* 15 (1): 117–131.

Wesch, Andreas. 1997. "El castellano hablado de Barcelona y el influjo del catalán: Esbozo de un programa de investigación [The spoken Spanish of Barcelona and the influence of Catalan: Outline of a research program]." *Anuario Galego de Filoloxía* 24: 287–312.

Wheeler, Max W. 1979. *Phonology of Catalan*. Oxford: Blackwell.

Young, Richard, and Robert Bayley. 1996. "VARBRUL Analysis for Second Language Acquisition Research." In *Second Language Acquisition and Linguistic Variation*, ed. by Robert Bayley, and Dennis Preston, 253–306. Philadelphia, PA: John Benjamins. DOI: 10.1075/sibil.10.11you