Journal of Phonetics 40 (2012) 639–654

Contents lists available at SciVerse ScienceDirect

Journal of Phonetics

journal homepage: www.elsevier.com/locate/phonetics

Voicing assimilation in Catalan two-consonant clusters

Daniel Recasens [a,b,*], Meritxell Mira [a,b]

[a] Departament de Filologia Catalana, Universitat Autònoma de Barcelona, Bellaterra, Barcelona 08193, Spain
[b] Laboratori de Fonètica, Institut d'Estudis Catalans, c/Carme, 47, Barcelona 08001, Spain

ARTICLE INFO

Article history:
Received 21 May 2011
Received in revised form
24 May 2012
Accepted 5 June 2012
Available online 7 July 2012

ABSTRACT

This paper reports electroglottographic (EGG) data for consonant sequences composed of a word final stop or fricative followed by a voiced consonant produced by eight speakers of a Romance language, i.e., Catalan, where these clusters undergo regressive voicing assimilation. Analysis results reveal considerable speaker- and consonant-dependent differences in the temporal period of vocal fold vibration during C1. In agreement with the degree of articulatory constraint (DAC) model of coarticulation, there appears to be a direct relationship between the extent to which consonants allow contextual voicing (voicing coarticulation resistance) and exert voicing effects on other consonants (voicing coarticulation aggressiveness) in a good number of cases; in other cases, however, this prediction does not hold, mainly in fricative+nasal, lateral sequences presumably due to the aerodynamic requirements involved. EGG and acoustic data for two-obstruent cluster pairs where C2 may be underlyingly voiced or voiceless but agrees in place and manner of articulation show that speakers may use not only the temporal extent of vocal fold vibration but also C1 and preceding vowel duration (as well as fricative noise intensity in clusters with C1=/s/) as voicing cues; in particular, segmental duration was found to stay more constant than vocal fold vibration across speakers. In view of this cooccurring relation, it is concluded that regressive voicing assimilation in Catalan may be signaled by vocal fold vibration and segmental duration and intensity acting interactively.

1. Introduction

This study is an investigation of voicing adaptation in heterosyllabic consonant clusters in Catalan. According to descriptive and theoretical accounts, in Catalan, syllable final obstruents (stops and fricatives, as well as affricates which will not be subjected to analysis in the present paper) are necessarily voiceless word finally before a pause (i.e., final devoicing), and assimilate in voicing to a following word initial consonant (Recasens, 1993). Final devoicing accounts for the alternation between voiced allophones intervocalically and voiceless ones prepausally, such as [β] and [p] of /b/ ([səˈβɛ] saber "to know", [sap] sap "he/she knows"), and [z] and [s] of /z/ ([bəˈzɛt] vazet "little glass", [bas] vas "glass"). The failure to find robust voicing phonetic attributes for word final obstruents before a pause in Catalan indicates that final obstruent devoicing operates completely in this language (Mascaró, 1987). Regressive voicing assimilation in clusters is considered to occur across a word boundary, in such a way that the voiceless word final realization of an underlyingly voiced or voiceless obstruent becomes voiced before a voiced consonant, e.g., [ən sab ˈmol] en sap molt "he/she knows a lot", [baz ˈbujt] vas buit "empty glass" (Bonet & Lloret, 1998). Another noteworthy fact about voicing in Catalan is an asymmetric behavior between word final stops and fricatives before a word initial vowel, i.e., fricatives but not stops exhibit a voiced realization in this position except for /f/ as discussed below ([baˈz amplə] vas ample "wide glass", [sap əˈʃɔ] sap això "he/she knows this").

The present investigation is concerned with the contribution of vocal fold vibration as well as other acoustic characteristics, i.e., segmental duration and intensity, to the phonetic implementation of the regressive voicing assimilation process in Catalan clusters composed of a word final obstruent followed by a word initial voiced C2. All throughout the paper, Catalan word final stops and fricatives will be transcribed with the voiceless symbol independently of their underlyingly voiced or voiceless status.

1.1. Voicing adaptation

1.1.1. VOT in simple stops

A first hypothesis to be tested is whether, as suggested by Westbury (1975), the degree of voicing adaptation in obstruent consonant clusters may be conditioned by the voicing characteristics of single consonants. If so, there should be a trend for languages where underlying voiced stops show a long voicing lead

* Corresponding author at: Departament de Filologia Catalana, Universitat Autònoma de Barcelona, Bellaterra, Barcelona 08193, Spain. Tel.: +34 3 4307952; fax: +34 3 5812782.
E-mail address: daniel.recasens@uab.es (D. Recasens).

0095-4470/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.wocn.2012.06.001

and underlying voiceless stops a moderate or no voicing lag to exhibit regressive voicing in clusters (Catalan and Romance languages in general, Russian). Thus, for example, /tɡ/ and /bk/ are, respectively, fully voiced and voiceless phonetically in Russian. On the other hand, languages where voiced stops are implemented through a short voicing lead or lag and voiceless stops are aspirated are expected to allow for mixed voicing adaptation scenarios (German, English). Thus, for the most part, English data reveal the presence of a voiced–voiceless realization of voiced+voiceless clusters, and of a voiceless–voiced or entirely voiceless realization of voiceless+voiced clusters (Westbury, 1979).

Several recent studies have shown that complete C1-to-C2 voicing adaptation fails to occur in many instances in languages and dialects exhibiting voiced stops with voicing lead, thus questioning Westbury's hypothesis regarding the relation between voicing in single consonants and voicing adaptation in clusters. Thus, French journalistic speech data (Hallé & Adda-Decker, 2007) show that regressive voicing in clusters is phonetically gradient. Also in Eastern Catalan, regressive voicing adaptation appears to vary with the consonant articulatory characteristics (Cuartero, 2001): while clusters with a voiceless C2 show absence of vocal fold vibration all throughout, C1 before a voiced C2 may be fully voiced or partially voiceless in obstruent+obstruent sequences such as /tɡ, sd/, or mostly or entirely voiceless in obstruent+sonorant sequences such as /kl, kn, sl, sn/.

1.1.2. The DAC model

Assuming that voicing adaptation in Catalan clusters applies to different degrees, a working hypothesis to be tested in the present investigation is whether the extent to which a given consonant induces voicing in an adjacent consonant increases the more it favours voicing in different environmental conditions (next to vowels, other consonants or a pause). This scenario is in accordance with the degree of articulatory constraint (DAC) model of coarticulation which has been formulated in order to account for data on lingual coarticulation so far (Recasens, Pallarès, & Fontdevila, 1997). According to the DAC model, lingual coarticulatory resistance for consonants to the effects of contextual vowels increases with the involvement of the tongue body in closure or constriction formation and is therefore greater for the alveolopalatal /ɲ/ than for the bilabial /p/ or the alveolar /n/. Moreover, in comparison to less resistant consonants, those which are more resistant are also more aggressive and, therefore, exert more prominent tongue body coarticulatory effects on the adjacent vowels (Recasens & Espinosa, 2009). Likewise, consonants allowing unimpeded airflow through the supraglottal cavities appear to be maximally resistant to devoicing effects induced by other consonants and at the same time are expected to cause other consonants to acquire voicing to the largest extent. Therefore, consonants which are more positively specified for a given articulatory property, i.e., voicing or tongue dorsum raising/fronting, are more resistant regarding variations induced in that property and also more aggressive.

1.1.2.1. Coarticulatory resistance. Regarding the voicing dimension, coarticulatory resistance for consonants is expected to vary with manner and place of articulation. Manner of articulation requirements account for why, in comparison to sonorants (nasals, laterals and approximants), obstruents (stops and fricatives) and the apical trill are less prone to allow continuous voicing. Nasals, laterals and approximants remain fully voiced considering that they are produced with relatively unimpeded continuous airflow through the oral or nasal cavities. As for their voicing adaptation behavior, sonorants do not undergo regressive devoicing assimilation before a heterosyllabic voiceless obstruent in Catalan and other Romance languages, and tend to exhibit vocal fold vibration all throughout in these circumstances (see, for example, voicing data for /l, ŋ/ before /t, s/ in Catalan in Cuartero, 2001). Continuous voicing is harder to maintain during the trill /r/ due to the high intraoral pressure level which is required to set the tongue into vibration and to facilitate the implementation of successive apical contacts (e.g., in Catalan, alveolar trill or trill-like realizations may be partly or fully voiceless utterance-finally and before a voiceless consonant; Recasens & Espinosa, 2007). Voicing may cease during closure for stops as the air volume stored within the occluded vocal tract causes the intraoral pressure to rise above the subglottal pressure level, and during fricatives due to the conflicting requirements between vocal fold approximation for voicing, on the one hand, and the need to allow much airflow through the glottis and to build up the necessary oral pressure for driving the noise source, on the other hand (Ohala & Solé, 2010; Westbury & Keating, 1986). Moreover, judging from literature reports, it seems that fricatives are more prone to devoice than stops in consonant clusters. Thus, in Dutch, voiceless+voiced clusters devoice the second consonant if it is a fricative rather than a stop (Slis, 1986).

Voicing degree for consonants may also be conditioned by place of articulation. Regarding stops, back articulations are less prone to be affected by voicing than front ones in line with differences in back cavity size and in the associated degrees of compliance of the vocal tract surface walls (Ohala & Riordan, 1979). Thus, in French, /p, t/ were found to assimilate to a following voiced consonant to a larger extent than /k/ while there was no preference among /b, d, ɡ/ to devoice before a voiceless consonant (Snoeren, Hallé, & Seguí, 2006). In parallel to stops, the devoicing of fricatives has been reported to affect /z/ rather than the more anterior articulations /v/ and /ð/ in English presumably in line with differences in vocal tract compliance, and also since voicing is more easily maintained during the shorter and weaker frication noise of (labio)dentals vs (palato)alveolars (Haggard, 1978; Pirello, Blumstein, & Kurowski, 1997; Stevens, Blumstein, Glicksman, Burton, & Kurowski, 1992). Data for other languages indicate, however, a trend towards less, not more voicing for the labiodental fricative than for lingual fricatives exhibiting more retracted places of articulation. Thus, in Russian, Hungarian and Czech, /v/ appears to be fully voiced intervocalically and after a voiceless consonant and devoices before a voiceless consonant but, contrary to stops and other fricatives, does not trigger voicing in a preceding consonant (Dvorak, 2010; Lulich, 2004; Markó, Gráczi, & Bóna, 2010). Also, unlike other fricatives, in Catalan, /f/ may not voice in intervocalic word final position ([baf uˈmit] <u>baf humit</u> "wet mist") and may resist regressive voicing assimilation in clusters with a voiced C2 (Recasens, 1993). A possible rationale for this voicing behavior may be the need for Catalan speakers to preserve the integrity of the labiodental fricative by enhancing its weak acoustic properties mostly in the light of the scarce number of Catalan words ending in /f/ many of which are of foreign origin (Mascaró & Rafel, 1990).

To summarize, a scale of degrees of voicing coarticulatory resistance for consonants may be formulated such that nasals, laterals and approximants are most resistant, and fricatives, stops and the apical trill are least resistant (and stops are more resistant than fricatives). Coarticulatory resistance for obstruents is also expected to increase with closure or constriction fronting, though not necessarily for /f/ which may be less resistant than /s/ and /ʃ/.

1.1.2.2. Coarticulatory aggressiveness. As predicted by the DAC model, if the degree of voicing adaptation in clusters is related to voicing degree in the triggering consonant, regressive voicing in Catalan should be more prone to occur as a function of sonorants than of obstruents and the apical trill, of stops vs fricatives, and of front vs back articulations though not necessarily in the case of fricatives.

Experimental evidence from the literature are by no means conclusive in this respect. Data for Dutch voiceless+voiced obstruent clusters reveal, as expected, that regressive voicing occurs more often in clusters where C2 is a stop than in those where it is a fricative (Slis, 1981). However, different results have been obtained for English, i.e., more, not less regressive voicing for /kz/ than for /kd/, which has been attributed to voiced fricatives being positively specified for [voice] and voiced stops for [-tense] (Jansen, 2004).

Another debatable case is regressive voicing of obstruents before sonorants. According to our approach, there should be maximal regressive voicing here since simple sonorants exhibit complete voicing. This prediction runs against the alternative hypothesis that sonorants should not trigger voicing. The line of argumentation in this case is that the fact that non-contrastive voicing for sonorants follows spontaneously from the low supraglottal pressure level associated with oral opening renders these consonants unspecified for the voicing feature (Jansen, 2004), or else more passively specified than voiced obstruents, which require specific articulatory adjustments such as voice onset time and pharyngeal expansion (Rice, 1993; Steriade, 1995). In agreement with this view, /s/ stays voiceless before /l, r/ in German (Beckman, Jessen, & Ringen, 2009), and the underlying voicing distinction for obstruents is maintained before sonorants but not before obstruents in Russian (Hall, 2007; Kulikov, 2011). There are reasons to believe that exceptions to the regressive voicing assimilation process (as well as cases of partial voicing during C1) before nasals and laterals are production-based to a large extent. Regarding obstruent+nasal clusters, a feasible explanation may be sought in the need to preserve the pressure difference across the oral constriction for intense turbulence for fricatives and the pressure buildup for the generation of a salient burst for stops, which could be impaired if velic lowering and voicing were both anticipated during C1 (Ohala & Solé, 2010). Also for obstruent+lateral sequences, the complete or partial lack of C1 voicing may be accounted for assuming that the anticipation of the oral constriction gesture for the lateral could endanger those same aerodynamic requirements by increasing the intraoral pressure level and perhaps by preventing the formation of a precise central constriction during the preceding fricative (see Beckman et al., 2009 for a similar interpretation for /z/ devoicing before /l, r/ in German), and by causing changes in vocal tract volume and compliance for a stop C1. The reason why sonorants may trigger regressive voicing in some languages but not in others could also be sought in specific language-specific production constraints for obstruents and/or sonorants in the clusters of interest.

A scenario worth exploring is that of /b, d, ɡ/ preceded by the fricative consonants /f, s, ʃ/. In Catalan, these clusters show C2 lenition to the approximants [β, ð, ɣ], except for the sequence /fb/, e.g., [vaz ˈβɔ] vas bo "good glass", [baf ˈbɔ] baf bo "good mist" (Recasens, 1993). In practice, however, C2 may also show a voiced stop realization in favorable conditions, e.g., at slow speech rates and in stressed position. Of much relevance to the present study is the interaction between lenition and voicing. Thus, given that approximants involve a lower intraoral pressure level and are shorter than stops, the former are expected to remain voiced and consequently to induce voicing in C1 to a larger extent than the latter.

The present paper will also investigate whether, in parallel to languages with no voicing lead (see Section 1.1.1), voicing adaptation in Catalan clusters may apply at the progressive level. These effects are prone to be purely coarticulatory and thus quite variable and of a limited temporal extent since progressive voicing assimilation is not expected to take place in this language. Progressive voicing adaptation has been reported to occur in other languages exhibiting voiced stops with voicing lead: in French clusters, while a voiceless C2 turns out to be hardly influenced by a preceding voiced consonant, a voiced C2 may exhibit a lower voicing degree after a voiceless consonant than after a voiced consonant although this voicing difference happens to be much smaller than the one occurring at the regressive level (Hallé & Adda-Decker, 2007). Moreover, crucially for the clusters under analysis in the present investigation and in agreement with predictions from the DAC model (see Section 1.1.2.1), fricatives ought to be more prone to trigger devoicing than stops in the following voiced consonant presumably since the amplitude of the devoicing gesture is usually larger for voiceless fricatives than for voiceless stops mostly if the latter are unaspirated (Hoole, 1999; Löfqvist & McGarr, 1987; Slis, 1981).

1.2. Interaction of voicing phonetic properties

Another central research issue to be addressed in the present investigation is whether voicing adaptation in clusters is implemented through not only vocal fold vibration but segmental duration and intensity as well. Data on consonant clusters reveal that obstruents are longer (and the preceding vowel shorter) if phonetically voiceless than voiced; thus, /k, ɡ, ʃ, ʒ/ in Hungarian have been reported to be shorter, and the preceding vowel longer, if occurring before a voiced vs voiceless consonant (Jansen, 2004). There may also be differences in the relative salience of the C1 and preceding vowel duration: French data indicate that differences in vowel duration remain more constant than differences in C1 duration with changes in the C2 voicing status, which suggests that the former parameter is perceptually more robust than the latter (Abdelli-Beruh, 2004). Moreover, these voicing-dependent vowel duration differences may be greater before fricatives than before stops (Laeufer, 1992). In principle, the C1 voicing distinction in consonant clusters ought to be also cued by a greater intensity of the fricative noise (Balise & Diehl 1994; Pirello et al., 1997) and by a greater intensity and a longer duration of the stop burst (Crystal & House, 1988; Zue, 1980) when the consonant is voiceless than when it is voiced. Several aerodynamic factors and production mechanisms, i.e., intraoral pressure level, airflow volume, degree of closure or constriction, as well as vocal fold abduction degree for fricatives (Kohler, 1984), appear to be responsible for the differences in segmental duration and intensity just referred to.

An open issue is the relative power of the phonetic properties and, more specifically, whether vocal fold vibration, segmental duration and intensity or all these factors combined may be considered to be the primary signaling cue of the voicing distinction. The fact that vocal fold vibration may play a secondary role in marking the underlying stop voicing distinction in languages without voicing lead (e.g., voiced stops, though not voiced fricatives, are regularly voiceless utterance initially and may be voiceless in intervocalic position in languages like German; Jessen, 2004, Beckman et al., 2009) has led scholars to advocate the view that the primary stop voicing cue should be [±fortis] or [±tense] instead of [±voiced] in these languages. Consistently with this view, data for clusters reveal that, compared to vocal fold vibration, C1 duration and airflow and vowel duration appear to be a more robust, less variable voicing characteristic for /zC/-/sC/ sequence pairs in English (Smith, 1997), and a similar finding has been reported for /fb/-/fp/ and /sd/-/st/ in Dutch (van Dommelen, 1983).

As for languages with voicing lead, the traditional view is that vocal fold vibration should be the primary voicing attribute

(Kohler, 1979). In support of this possibility, French data reveal that vocal fold vibration is more salient than segmental duration (Snoeren et al., 2006), and Russian data that segmental duration operates in fricative+stop clusters (see above) but not in stop+fricative clusters where the voicing distinction is signaled by vocal fold vibration exclusively (Burton & Robblee, 1997). Contrary to the prevailing belief, however, it has even been proposed that voicing in Romance languages like Spanish is cued primarily not by vocal fold vibration but by stop closure and preceding vowel duration (Martínez Celdrán & Fernández Planas, 2007). This proposal is based on production and perceptual evidence as well as on the observation that unaspirated voiceless stops may lenite and become voiced in contexts favoring consonant reduction (e.g., in intervocalic word medial position). The present study will look into the possibility that speakers of Catalan may have learnt how to keep the voicing distinction in two-consonant clusters by controlling other articulatory dimensions besides vocal fold vibration which are more directly related to oral pressure and airflow volume such as C1 duration and intensity and preceding vowel duration.

1.3. Summary of research goals

A first goal of the present investigation is to determine the extent to which regressive voicing adaptation in Catalan obstruent+voiced consonant sequences conforms to the principles of the DAC model. We expect consonants to exhibit different degrees of resistance to variations in voicing degree depending on the specific articulatory and aerodynamic factors involved in their production, and coarticulatory aggressiveness (the extent to which consonants induce voicing in other consonants) to increase with voicing coarticulation resistance. The paper also evaluates the special voicing behavior of /f/, the presence of progressive devoicing, and the relationship between regressive voicing assimilation and VOT, the hypothesis being that speakers showing higher negative VOT values in underlying voiced stops ought to favor regressive voicing assimilation to the largest extent. The power of phonetic attributes other than vocal fold vibration in marking the voicing contrast will also be investigated for pairs of clusters composed of an obstruent C1 and a voiced/voiceless C2. The expected behavior in this case is for the C1 closure and burst or frication period to be shorter, and for the preceding vowel to be longer, in clusters with a voiced vs voiceless C2, and for C1 frication to be more intense in clusters with a voiceless vs voiced C2. The extent to which these duration and intensity characteristics become prevailing phonetic characteristics of voicing in the complete or partial absence of vocal fold vibration during C1 will also be investigated. Finally, speaker-dependent differences in the implementation of voicing both regarding degree of vocal fold vibration and the other parameters will be evaluated within the framework of the DAC model.

The onset and offset of vocal fold vibration will be identified on glottal waveforms using electroglottography (EGG) which provides direct information on vocal fold opening and closing during vocal fold vibration by measuring the change in electrical impedance across the throat (Rothenberg & Mahshie, 1988). There are problems with other voicing analysis methods such as the inspection of the voicing bar on spectrograms which do not apply to EGG and therefore render the latter method more suitable than the former: the voicing bar may correspond to glottal oscillations rather than to true glottal pulses (see Section 2.2.2), and true glottal pulses may not be visible on spectrograms due to background noise or the low intensity level of the acoustic signal. EGG has been used for the analysis of voicing adaptation in consonant clusters in several publications referred to in this paper (Cuartero, 2001; Slis, 1981; Smith, 1997).

2. Method

2.1. Experimental paradigm

Two-consonant sequences and several single consonants were recorded in intervocalic position in the five- to seven-syllable long Catalan meaningful sentences listed in the Appendix. They were placed next to a (mid) open vowel practically in all sentences, and were followed by a vowel carrying lexical and sentence stress in all cases except for sentences 56–61 where the vowel in question carried lexical stress only. In Catalan, /t, d/ are dental stops, /l/ is a dark alveolar lateral, /r/ is an alveolar trill and /ʎ/ is an alveolo-palatal lateral. The recording material was organized into the following four sentence groups:

(a) All possible combinations of word final C1=/p, t, k, f, s, ʃ/ and word initial C2=/b, d, g, m, n , l, z, r, ʎ, j/ for the analysis of regressive voicing assimilation (sentences 1–55 in the Appendix), except for /pb, td, kg, sz, ʃz/ which are realized systematically as a long consonant ([bb, dd, gg, (z)z, (z^j)z^j]). The presence of an approximant or stop realization of /b, d, g/ after /f, s, ʃ/ allowed looking into the interaction between lenition during C2 and voicing during the two consonants of the cluster. Several sequences were excluded from analysis: long consonant realizations originated more or less occasionally from /tb, pm, tm, tn, tl, sr, tr, tʎ/ ([bb, mm, nn, ll, (r)r, ʎʎ]) through regressive assimilation of place and/or manner of articulation as assessed auditorily and through inspection of spectrographic displays by the first paper author; other sequences derived through regressive manner assimilation where C1 was realized neither as a stop nor as a fricative (/pn, kn/ > [mn, ŋn]).
(b) The stops /p, t, k, b, d, g/ in postpausal CV sequences with the vowel /a/ in order to measure voiced onset time, and to correlate negative VOT values for /b, d, g/ with the extent to which C1 undergoes voicing assimilation before those three consonants in clusters (sentences 56–61 in the Appendix).
(c) The fricatives /f, s, ʃ/ in intervocalic word final position so as to check the extent to which their voicing degree matches the voicing degree for fricatives in consonant clusters (sentences 62–64 in the Appendix). It should be recalled at this point that intervocalic word final fricatives are expected to be phonetically voiced in Catalan (see Section 1).
(d) Ten cluster pairs with a given word final obstruent followed by obstruents differing between each other in underlying voicing but not in place or manner of articulation, e.g., /pt-pd/ in the sentences no queda ca**p t**alp "there isn't any mole left" and no queda ca**p d**au "there isn't any dice left" (sentences 65–84 in the Appendix). They were used for the analysis of the relative power of several voicing attributes, i.e., vocal fold vibration during C1, the duration of C1 and V (the vowel preceding the cluster), the intensity of the C1 frication noise, and the duration and frequency of occurrence of the C1 stop burst. These cluster pairs were classified into three groups, i.e., stop+stop (/pt-pd/, /pk-pg/, /kp-kb/, /kt-kd/), stop+fricative (/ps-pz/, /ts-tz/, /ks-kz/) and fricative+stop (/sp-sb/, /st-sd/, /sk-sg/). The clusters pairs /tp-tb/ and /tk-tg/ could not be included since C1 may undergo regressive place assimilation in this case.

Acoustic and electroglottographic data were recorded simultaneously by eight native Catalan speakers, i.e., five women (EV, MA, PE, LO, and VA) and three men (SO, MO, and DR) of about 25–55 years of age, using the multichannel Kay Pentax system. These informants come from different areas of Catalonia: six of them speak the Eastern Catalan dialect and were born in urban Barcelona (SO, PE) and in other towns and villages (MO, Banyoles; LO, Montblanc; DR, Tarragona; VA, Cadaqués); the remaining two subjects speak Western Catalan and were born in the Baix Urgell region (EV, MA). No differences in degree of regressive voicing assimilation were expected to occur as a function of the (sub)dialectal variety of Catalan spoken in Catalonia. All sentences were read eight to ten times at the speakers' normal speech rate, and seven of these repetitions were chosen for analysis. The acoustic data were acquired with a Shure SM58-LCE microphone, and the EGG data with an EGG-2 glottograph from Glottal Enterprises by placing two surface electrodes onto the speaker's neck. Both signals were acquired at 44,100 Hz, and downsampled to 500 Hz the EGG signal, and to 11,025 Hz the acoustic signal for better inspection of the spectral events occurring within the 0–5.5 kHz range on spectrographic displays. The EGG signal was smoothed using three steps depending on the degree of noisiness of the glottal signal and analyzed using the MatLab script Peakdet 2 (Abadal & Recasens, 2009).

2.2. Measurement criteria

2.2.1. Segmentation

Onsets and offsets were estimated for V, C1 and C2 in clusters and for single consonants, based on visual inspection of simultaneous spectrographic and waveform displays using CSL (Computerized Speech Lab) from Kay Pentax.

Phonetic segments were delimited by the edges of a period of high intensity formant structure for vowels, of low intensity formant structure for nasals, laterals and approximants, of acoustic closure with no available formants for stops, and of a high frequency frication noise for fricatives. Based on visual inspection of spectrographic displays, C2=/b, d, ɡ/ after a fricative were classified as [β, ð, ɣ] (approximants) if exhibiting weak formants occasionally with some frication noise overimposed, or as [b, d, ɡ] (voiced stops) if showing no formant structure and a burst. The alveolar trill was generally identified by the presence of two or more short closures separated by short opening phases (see Fig. 1); if realized sometimes as an approximant or a fricative, the onset and offset of /r/ were determined applying the same criteria for /b, d, ɡ/ above.

Stop bursts were considered not to be part of the stop consonant and could be absent for C1 in several stop+stop cluster repetitions, which were excluded from analysis. Whenever the velar stop burst exhibited a double spike, the stop burst was taken to start at the first spike if the burst including this spike was about 30 ms long, which is the regular duration of the unaspirated voiceless velar stop burst in Catalan (Recasens, 1986); if the duration of the velar stop burst exceeded 30 ms, the burst was taken to extend from the second spike to vowel onset (in these circumstances, the first spike was considered to be generated as the slow and massive tongue dorsum slides over the palate surface prior to release). A voiceless period occurred often at the boundary between a fricative C1 and a nasal C2, and, less often, in stop+nasal and in stop, fricative+lateral sequences as well (see Section 1.1.2.2). Following previous accounts on fricative+nasal sequences (English, Docherty, 1992; Dutch, van Dommelen, 1983), the period in question was considered to be part of C2 as a general rule and, as exemplified in Fig. 2, could exhibit weak or no nasal formants.

2.2.2. Detection and analysis of voicing

The present study establishes differences in degree of consonant voicing based on a statistical evaluation of voicing percentages over consonant duration. Percentages have been used as a voicing measure in the literature (Docherty 1992; Haggard, 1978; Myers, 2002; Smith, 1997; Snoeren et al., 2006), and

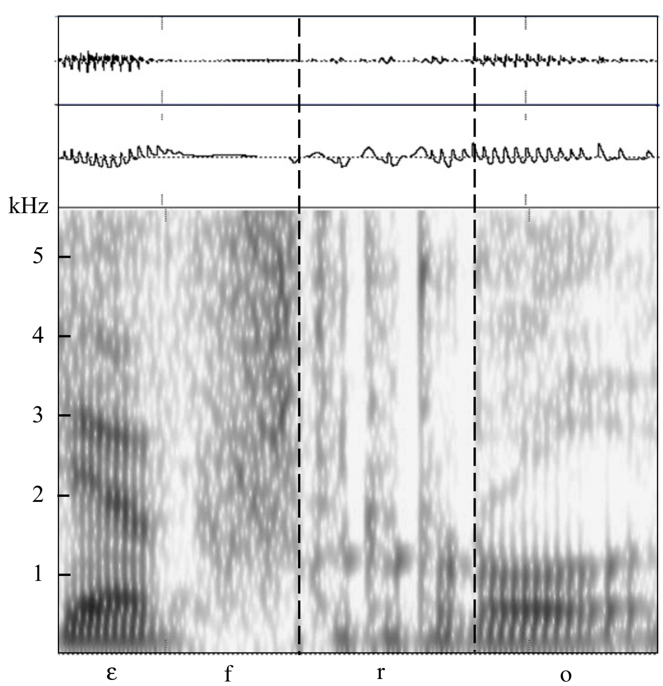

Fig. 1. Acoustic waveform (top), glottal waveform (middle) and spectrogram (bottom) for the sequence /ɛfro/ occurring in the sentence 'a la cuina hi ha un xef ros' ("there is a blond chef at the kitchen"). Vertical lines have been inserted at the onset and offset of the alveolar trill. Data correspond to speaker MA.

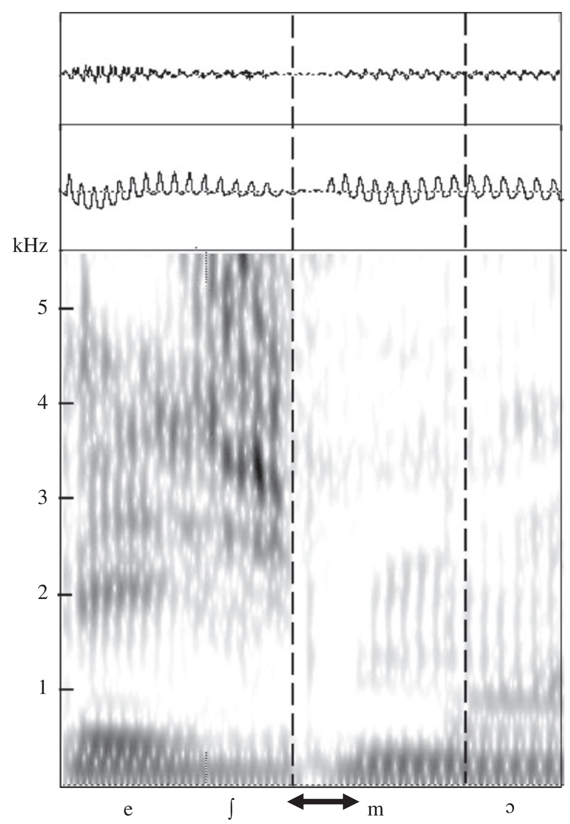

Fig. 2. Acoustic waveform (top), glottal waveform (middle) and spectrogram (bottom) for the sequence /eʃmɔ/ occurring in the sentence 'no l'agafis, el feix moll' ("do not take the wet bundle"). Vertical lines have been inserted at the onset and offset of /m/, and a double arrow indicates the presence of a voiceless period devoid of nasal formants at the onset of the nasal murmur. Data correspond to speaker DR.

represent a way to normalize the voicing data across variations in segmental duration associated with speaker and speech rate.

Peakdet 2 was used for identifying the onset and offset of voicing and, therefore, for measuring the duration of the voicing phase(s) during a given consonant. The time at which a voicing pitch pulse occurs is identified by Peakdet 2 at the positive peak of the first derivative of the glottal waveform (DEGG) which corresponds to the glottal closing instant. This peak picking procedure was applied setting a threshold detection at 25% of the DEGG positive maximum which is slightly below other threshold values proposed in the literature (Rothenberg & Mahshie, 1988). In order to account for the presence of double DEGG peaks, the peak picking procedure was carried out using the barycentre method which weighs the two peaks and takes a temporal point close to the highest peak (Henrich, d'Alessandro, Doval, & Castellengo, 2004; Mazaudon & Michaud, 2008).

Both in simple consonants and in clusters, continuous voicing as indicated by the Peakdet 2 program could be present or absent all throughout, or else be interrupted for a shorter or longer period of time after an initial period of voicing following immediately the preceding vowel. Much less often, there were several alternating periods of voicing and voicelessness (a clear exception was the apical trill which was produced with voiceless contacts and voiced opening periods most of the time; see Fig. 1). Whenever C1 voicing was continuous with voicing in the preceding vowel, the onset and offset of vocal fold vibration for C1 and C2 were taken to occur at their acoustic onset and offset, respectively. For periods where voicing was discontinuous, labeling was carried out at the first glottal pulse of the voicing period following a period of voicelessness as determined by the first DEGG peak, and at the closing state of the last glottal pulse of a voicing period as determined by the last positive DEGG peak. Two or more consecutive pulses had to be present for them to be attributed to a voicing period; therefore, isolated glottal pulses surrounded by periods of voicelessness were assigned to a voiceless period. A special case occurred when well-defined glottal pulses were replaced by quasi-periodic low amplitude glottal oscillations during which the folds make contact presumably at the anterior part of the glottis (Mazaudon & Michaud, 2008). These oscillations were not treated as true glottal vibrations since they did not exhibit a DEGG closing peak. They were visible towards the offset of a continuous period of voicing, and, as exemplified by Fig. 3, for lingual fricatives that were produced with much airflow passing through the glottis and a high intraoral pressure level (see also Jesus & Shadle, 2002; Pinho, Jesus, & Barney, 2009).

An issue needs to be addressed at this point, i.e., the extent to which voicing during C1 in clusters composed of an obstruent followed by a voiced consonant can be safely attributed to C2 or should be associated fully or in part to the vowel preceding the cluster. The latter option is suggested by the fact that some voicing associated with the preceding vowel (i.e., vowel voicing lag) may occur at the onset of a postvocalic voiceless stop or fricative whether appearing by itself in intervocalic position or acting as C1 in clusters with a voiceless C2. In order to investigate this issue, voicing lag was measured in several voiceless clusters appearing in the Appendix (sentences 56–61:/tk/ 'soldat curd', /fk/ 'xef curd', /ʃk/ 'peix car', /kt/ 'sac tort', /pk/ 'catxap curt', /sk/ 'envàs car') after identifying the last glottal pulse occurring after vowel offset. The mean value across clusters and speakers was 8.8 ms (sd=6.5, range=0–32.4 ms), which did not exceed 20% of the C1 duration with the exception of the voiceless clusters with a stop C1 for speaker VA whose voicing lag ranged between 20% and 55%. Overall, the value of interest appears to be low compared to the 50 ms threshold reported for Dutch (Slis, 1981), which means that voicing for C1 in Catalan clusters with a voiced C2 can be safely attributed to C2 rather than to the vowel preceding the cluster. Fig. 4 exemplifies the presence of vowel-dependent voicing into C1 in the case of the cluster /pt/.

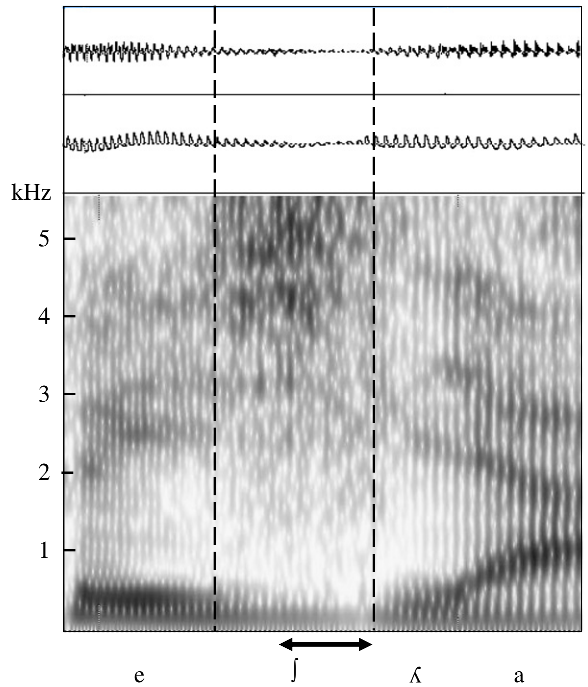

Fig. 3. Acoustic waveform (top), glottal waveform (middle) and spectrogram (bottom) for the sequence /eʃʎa/ included in the sentence 'necessitem un feix llarg' ("we need a long bundle"). Vertical lines have been inserted at the onset and offset of /ʃ/, and a double arrow indicates the presence of quasi-periodic low amplitude glottal oscillations which were not considered to be true glottal vibrations. Data correspond to speaker MA.

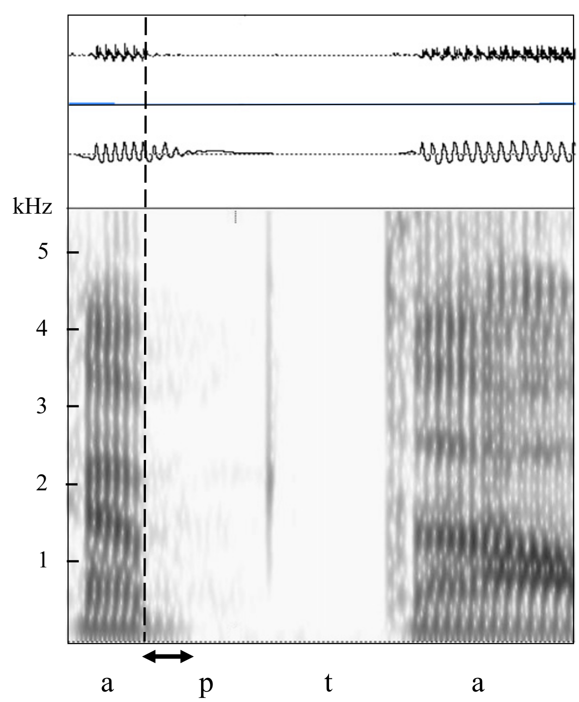

Fig. 4. Acoustic waveform (top), glottal waveform (middle) and spectrogram (bottom) for the sequence /apta/ included in the sentence 'no queda cap talp' ("there isn't any mole left"). A vertical line has been inserted at /p/ closure onset and a double arrow indicates the presence of vowel voicing into closure. Data correspond to speaker DR.

VOT measurements were carried out after identifying the first glottal pulse appearing before and after the stop burst. VOT was taken to be the temporal difference between the stop burst and the onset of voicing, and could be negative or positive depending on whether voicing onset occurred before or after the stop burst.

2.2.3. Evaluation of intersegmental voicing adaptation

In order to estimate the degree to which C1 assimilates in voicing to C2, i.e., the degree of C1 resistance to the C2 voicing effects, voicing percentages at C1 were averaged for each C1 across all C2 conditions (e.g., mean voicing percentages at C1 for C1=/f/ across C2=/b, d, ɡ, m, n, l, z, r, ʎ, j/). On the other hand, the degree to which C2 undergoes devoicing as a function of C1, i.e., the degree of C2 resistance to the C1-dependent devoicing effects, was computed by averaging the voicing percentages at C2 for each C2 across all six C1 consonant conditions (e.g., mean voicing percentages at C2 for C2=/z/ across C1=/p, t, k, f, s, ʃ/).

The degree to which C2 triggers voicing in C1, i.e., C2 aggressiveness, and C1 causes C2 to devoice, i.e., C1 aggressiveness, were evaluated as follows: C2 aggressiveness was calculated by averaging the voicing percentages at C1 for each C2 across all C1 conditions (e.g., mean voicing percentages at C1 for C2=/z/ across C1=/p, t, k, f, s, ʃ/); C1 aggressiveness was estimated by averaging the voicing percentages at C2 for each C1 across all C2 conditions (e.g., mean voicing percentages at C2 for C1=/f/ across C2= /b, d, ɡ, m, n, l, z, r, ʎ, j/).

2.2.4. Other phonetic characteristics

Other phonetic attributes of voicing besides vocal fold vibration and segmental duration were subjected to analysis in the case of the cluster pairs listed in Section 2.1(d). This was so for the duration and frequency of occurrence of the C1 burst in stop+stop sequences but not in stop+fricative sequences where the stop burst could not be easily distinguished from the /s, z/ frication noise in most cases. The absolute energy level of the frication noise was also measured at C1 midpoint in /s/+stop sequences on 10 ms window energy profiles using the energy display of the CSL system of Kay Pentax. Energy values in dB are obtained by multiplying intensity by duration (Dorman, Studdert-Kennedy, & Raphael, 1977). Downsampling was believed not to affect the fricative noise energy measurements since the spectral peak for /s/ in Catalan occurs at about 3500–4000 Hz (Recasens & Espinosa, 2006). Moreover, in order to normalize the absolute energy differences across speakers and to make sure that differences in /s/ intensity between the voiced and voiceless C2 conditions were not due to differences in syllable prominence, relative energy values were also calculated for each sequence token by dividing the absolute energy value at the midpoint of the fricative by that at the midpoint of the following vowel (Cho, Jun, & Ladefoged, 2002).

2.3. Statistical analysis

Several univariate ANOVAs were performed on the speakers' mean values with speaker as a random factor and the following dependent variables:

(a) Voicing percentages at C1 and at C2 for the clusters referred to in Section 2.1(a) with the analysis factors 'C1' (levels 'p, t, k, f, s, ʃ') and 'C2' (levels 'b, d, g, m, n, l, z, r, ʎ, j').
(b) Fricative voicing percentages for the VCV sequences referred to in Section 2.1(c) with the factor 'fricative' (levels 'f, s, ʃ').
(c) C1 and C2 voicing percentages and V, C1 and C2 duration values for the clusters referred to in Section 2.1(d) with the factor 'C2' (levels 'voiced', 'voiceless'). Separate ANOVAs were performed for the stop+stop, stop+fricative and fricative+stop sequences.
(d) Frication noise energy values for the fricative+stop sequences referred to in Section 2.1(d), and data on C1 burst duration and frequency of occurrence for the stop+stop sequences in the same section. The analysis factor was 'C2' (levels 'voiced', 'voiceless') in all cases.

Pairwise comparisons using the Bonferroni correction were carried out in order to uncover significant differences among levels of those analysis factors which yielded a main effect. Significant two-factor interactions were also analyzed statistically for each level of a significant independent variable. In all statistical tests, the significance threshold was set at $p=0.05$.

3. Results

3.1. Voicing in consonant clusters with a voiced C2

An analysis of the voicing adaptation patterns in clusters with a voiced C2 allows studying the degree of coarticulatory resistance and aggressiveness for the two consonants in the cluster. This section deals with the degree of resistance to voicing assimilation for C1 and C2 (Section 3.1.1), and with the extent to which the regressive and progressive voicing effects conform to differences in coarticulatory resistance among the triggering consonants (Section 3.1.2).

3.1.1. Coarticulatory resistance

(a) Consonants occupying the C2 position in clusters differed significantly in voicing ($F(9,63.37)=22.45$, $p<0.001$). To the extent that these voicing percentages have been obtained across all six C1 conditions, they may be said to reflect the degree to which a given C2 resists the coarticulatory influence of C1 and therefore the degree of C2 voicing coarticulation resistance. According to results from pairwise comparisons and as revealed by Fig. 5 (right graph), these consonant-dependent voicing percentages decreased in the progression the approximant /j/, the laterals /l, ʎ/, the nasals /m, n/ (84.9–94.7%) > the fricative /z/ (69.5%) > the stops /b, d, ɡ/, the trill /r/ (50.6–64.6%). C2 voicing percentages plotted along the vertical axis of the graphs in Fig. 6 reveal that the approximant, the laterals and the nasals (filled symbols) exhibit more voicing than the obstruents and the trill (empty symbols) for all individual speakers as a general rule, and that there is more voicing for /z/ than for stops and /r/ for about five subjects.

Moving back to Fig. 5, stop voicing at the C2 site was found to decrease in the progression labial /b/ > dental /d/ > velar /ɡ/. The degree of voicing for C2=/b, d, ɡ/ after /f, s, ʃ/ was 20–55% greater when the consonant was realized as an approximant (or much less often as a fricative) vs a stop; mean voicing percentages across clusters were 75.8% for [β, ð, ɣ] (sd=14.5) and 39.7% for [b, d, ɡ] (sd=15.1). Similar voicing percentages hold for the /sC/ clusters referred to in (d) in the Appendix, i.e., 81.5% for [β, ð, ɣ] (sd=18.9) and 48.2% for [b, d, ɡ] (sd=30.7).

(b) As for C1, a main consonant effect in voicing degree ($F(5,35.16)=7.14$, $p<0.001$) was associated with lower values for the fricatives /f, s, ʃ/ (39.4%) than for the stops /p, t, k/ (58.7%). Pairwise comparisons revealed the presence of highly significant place-dependent differences for /p, t/ > /k/ among stops and for /ʃ/ > /f/ among fricatives (/s/ did not differ significantly from either /f/ or /ʃ/). Fig. 7 (left graph) shows indeed lower mean voicing values for the velar (50.7%) than for the labial and the dental (61.7% and 63.7%, respectively) among stops, and for /f/ (32.7%) than for /s/ (41.1%) and /ʃ/ (45.1%).

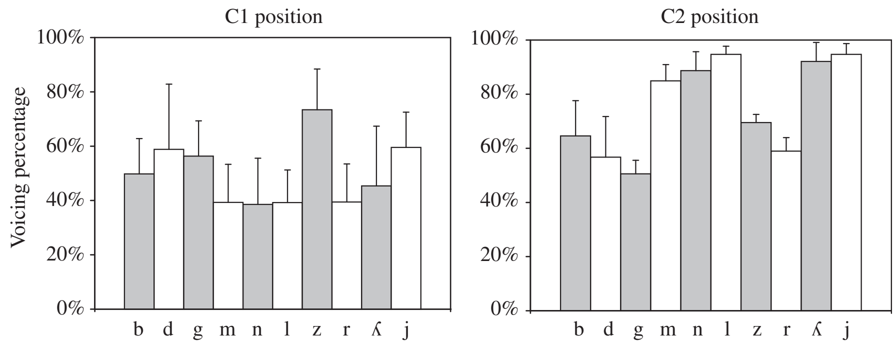

Fig. 5. Voicing percentages as a function of C2=/b, d, g, m, n, l, z, r, ʎ, j/ averaged across all C1 conditions and speakers both at the C1 position (left) and at the C2 position (right). The right graph is an indicator of C2 coarticulatory resistance and the left graph an indicator of C2-dependent coarticulatory aggressiveness. Error bars correspond to one standard deviation of the voicing means.

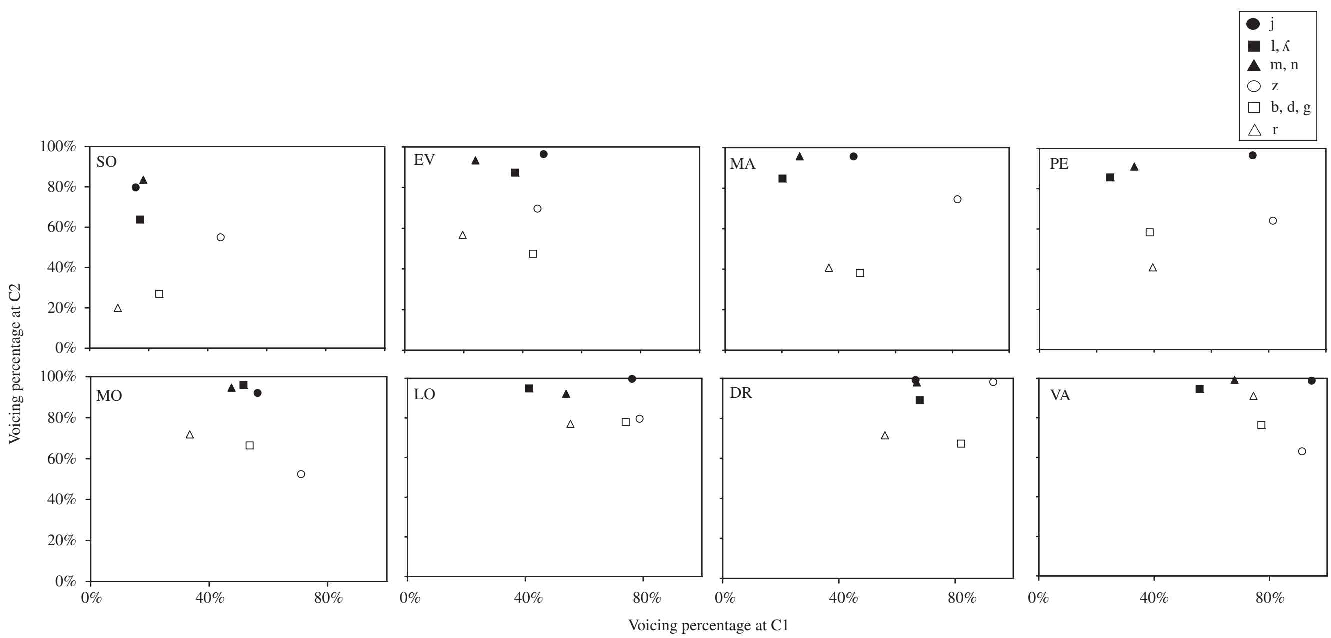

Fig. 6. C2-dependent voicing effects for the individual speakers SO, EV, MA, PE, MO, LO, DR and VA. In the graphs, voicing percentages for C2 at the C2 position (vertical axis) have been plotted against C2 voicing percentages averaged across C1 conditions at the C1 position (horizontal axis). Consonants are assigned different symbol types depending on voicing degree, i.e., filled symbols (maximal voicing), empty symbols (minimal voicing).

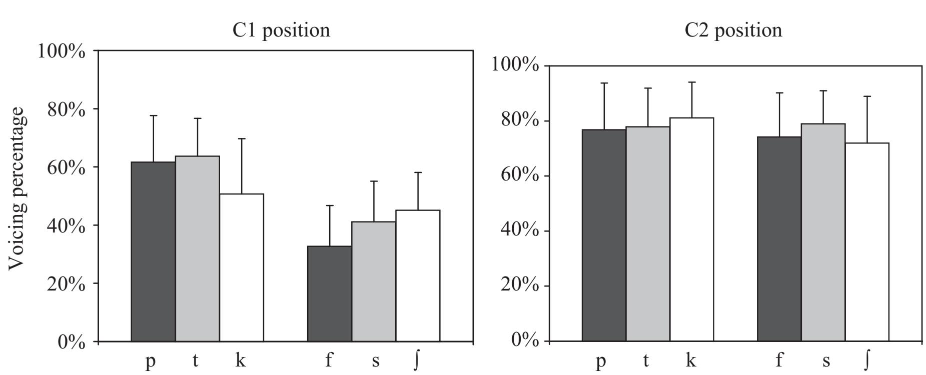

Fig. 7. Voicing percentages as a function of C1=/p, t, k, f, s, ʃ/ averaged across C2 conditions and speakers both at C1 (left) and at C2 (right). The left graph is an indicator of C1 coarticulatory resistance and the right graph an indicator of C1-dependent coarticulatory aggressiveness. Error bars correspond to one standard deviation of the voicing means.

These data indicate that regressive voicing in clusters, as implemented by vocal fold vibration during C1, is by no means complete. Moreover, they support to a large extent the consonant-dependent differences in coarticulatory resistance pointed out in Section 1.1.2.1. The voicing data for C1 indicate that fricatives are less resistant to changes in voicing degree, i.e., they may show less voicing during their production than stops, while these data and those for C2 show that sonorants (the nasals, the laterals and the approximants [j, β, ð, ɣ]) are more resistant to voicing variations than obstruents and that the apical trill is most prone to devoice. The only apparent exception is the alveolar fricative which shows less voicing in C1 position than when voiced underlyingly in C2 position. The expected place of articulation effect was found to hold for stops, while /f/ exhibited less voicing than the more retracted lingual fricatives /s, ʃ/.

Place-dependent differences in voicing degree among fricatives in C1 position in consonant clusters were consistent with those among simple fricatives (Section 2.1(c)). Indeed, intervocalic word final fricatives exhibited significant voicing differences as a function of place of articulation ($F(2,14)=9.55$, $p<0.01$), which turned out to be related to half as much voicing for /f/ (41% of the overall consonant duration, sd=36.2) as compared to /s/ (82.7%, sd=25.1) and /ʃ/ (82.0%, sd=26.6). Moreover, while similar voicing degrees were obtained for /f/ in intervocalic position and in C1 position in clusters, the voicing value for intervocalic /s/ and /ʃ/ was comparable to that for C2=/z/ and exceeded that for C1=/s, ʃ/ in clusters.

Speakers were found to differ significantly among themselves in C1 voicing degree ($F(7,39.97)=11.52$, $p<0.001$), i.e., C1 voicing percentages ranged between 20% and 90% depending on the speaker taken into consideration. Moreover, these speaker-dependent differences could be dialect-dependent: subjects speaking the Eastern Catalan dialect from urban Barcelona (SO, PE) or Western Catalan (EV, MA) exhibited relatively low voicing degrees, while speakers from Eastern Catalan areas other than Barcelona (MO, LO and, even more so, DR and VA) showed higher degrees of voicing (see Fig. 8, left graph). In addition to a main speaker effect, there was a significant speaker x C1 interaction ($F(35,271)=4.56$, $p<0.001$) which was related to manner-dependent voicing differences, i.e., six speakers showed less voicing for fricatives than for stops, speaker SO exhibited a very low voicing degree for all consonants, and speaker EV devoiced stops to a larger extent than fricatives. Speaker-dependent differences were obtained for intervocalic word final fricatives as well. Speakers differed mostly regarding the degree of voicing for /f/ which could be below 25% (EV, MA, PE, DR), about 50% (MO, LO) or above 80% (SO, VA). A significant consonant x speaker interaction ($F(4,144)=9.80$, $p<0.001$) was related to significant differences in voicing degree between the labiodental and the two lingual fricatives for all speakers except for SO and VA (who showed high voicing percentages for all three consonants) and for MA (who exhibited lower voicing percentages for /s, ʃ/ than all other subjects, i.e., about 40%).

3.1.2. Coarticulatory aggressiveness

(a) Voicing percentages during C1 varied depending on the following consonant ($F(9.63,74)=10.68$, $p<0.001$). Both for the cross-speaker and the individual speakers' data, C1 voicing degree decreased with C2 in the progression /z/ (73.4%) > /b, d, g, j/ (49.8–59.5%) > /m, n, l, r, ʎ/ (38.6–45.4%) (see Fig. 5, left graph). A comparison between C2-dependent differences in voicing at C1 and at C2 (compare the left and right graphs of the same figure) reveals that there is no necessary relationship between voicing degree at the two consonant locations and, therefore, between voicing coarticulatory resistance and voicing coarticulatory aggressiveness. In agreement with the initial prediction, the trill /r/ is prone to exhibit a low voicing percentage in C2 position (59%) and to cause C1 to exibit a small degree of voicing (39.5%), and the stops /b, d, g/ and the fricative /z/ exhibit similar voicing degrees during C1 (/b, d, g/ 49.8–58.8%, /z/ 73.4%) and C2 (/b, d, g/ 50.6–64.6%, /z/ 69.5%). Contrary to the hypothesis that coarticulatory resistance and aggressiveness should be positively related, however, nasals and laterals are mostly voiced (84.9–94.7%) while not inducing much voicing in the preceding consonant (38.6–45.4%); moreover, place-dependent differences in degree of voicing for stop consonants (labial > dental > velar) are not traceable during C1.

In order to study the extent to which individual speakers conform to this general voicing adaptation scenario, the graphs in Fig. 6 plot for each individual subject the voicing percentages for each C2 at the C2 site (vertical axis) against the voicing percentages for each C2 averaged across all C1 conditions at the C1 site (horizontal axis). If there was a positive correlation between voicing at the two consonant sites conforming to the prediction of the DAC model, the black symbols ought to appear towards the top right corner of the graphs since they correspond to the consonant exhibiting maximal voicing during C2. In parallel to the cross-speaker voicing percentages, a positive relationship between voicing resistance and aggressiveness occurs for stops, /z/ and /r/, i.e., the amount of voicing varies in the progression /z/ > stops, /r/ at both consonant sites, in the case of speakers SO, MA, PE, DR and to a lesser extent EV, but does not hold for the nasals and the laterals for any of the eight speakers subjected to analysis.

While C2=/b, d, g/ in clusters with a fricative C1 were more considerably voiced when realized as approximants than as stops (see Section 3.1.1), variations in C1 voicing degree as a function of the C2 manner of articulation in these clusters were found to occur only for /sd/ and /ʃd/. Indeed, the C1 voicing percentages for these two-consonant sequences increased by 23.5% and 32% when C2 was realized as an approximant (or occasionally as a fricative) as opposed to a stop. As for the /sC/ clusters in Section 2.1(d), higher C1 voicing percentages before approximant vs stop realizations were found to hold for /sb/ (14.6% difference) and /sg/ (13.2% difference).

(b) The progressive influence of C1 on the C2 voicing percentages did not achieve significance. There was however a

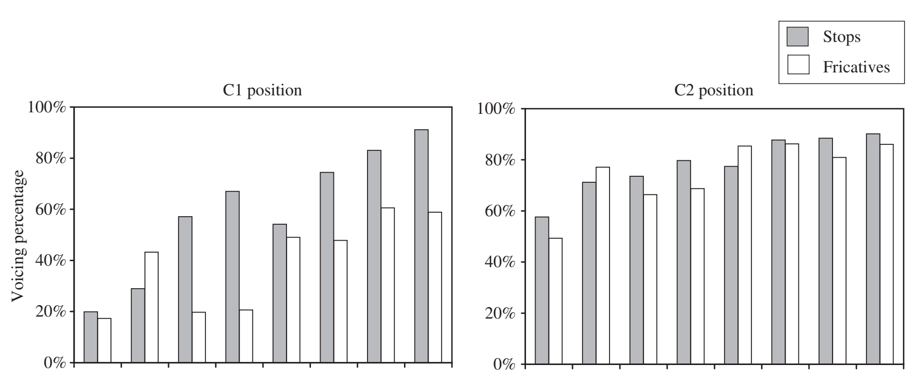

Fig. 8. Voicing percentages as a function of a stop C1 (/p, t, k/) and a fricative C1 (/f, s, ʃ/) averaged across C2 conditions at the C1 position (left) and at the C2 position (right) for all speakers.

significant C1 × C2 interaction ($F(40,271)=2.52$, $p<0.001$) which was associated mostly with a higher voicing degree for a stop C1 than for a fricative C1 at C2 when C2 was a voiced stop, i.e., for /t/ > /ʃ/ and /s/ > /f, ʃ/ at C2 = /b/, for /p, k/ > /ʃ/ and /k/ > /f, s/ at C2 = /d/, and for /p/ > /s, ʃ/ at C2 = /g/ (see Fig. 7, right graph). This finding is indicative of some progressive voicing adaptation in clusters and of a positive relationship between coarticulatory resistance and aggressiveness, i.e., C1-to-C2 devoicing effects are most prominent for those consonants allowing less voicing, i.e., fricatives vs stops. This trend holds for the individuals speakers as well. Thus, the right graph of Fig. 8 reveals the presence of slightly higher C2 voicing percentages after stops than after fricatives for all speakers with the exception of EV (who also shows more voicing for fricatives than for stops at the C1 position) and MO.

3.1.3. *Regressive voicing in clusters and VOT*

All speakers were found to conform to the voicing lead pattern and thus, showed a long voicing lead for underlying voiced stops and a moderate or no voicing lag for underlying voiceless stops. Mean VOT values for utterance initial stops across speakers (Section 2.1(b)) were negative for the underlying voiced stops and slightly positive for the underlying voiceless stops. Moreover, in agreement with data for other languages (Lisker & Abramson, 1964), negative VOT values for voiced stops were shorter for /g/ (−11.7 ms, sd = 25.7) than for /b, d/ (−53.1 ms, sd = 17.7; −57.8 ms, sd = 23.1), and positive VOT values for voiceless stops decreased in the progression /k/ (33 ms, sd = 4.6) > /t/ (23 ms, sd = 4.2) > /p/ (16.5 ms, sd = 4.3). The VOT value for /g/ was slightly positive for some speakers (SO, DR) and around 0 ms for other speakers (EV, PE). Correlation analyses between the C1 voicing percentages in clusters with C2 = /b, d, g/ and the VOT values for postpausal /b, d, g/ yielded a very low r value which indicates that speakers showing higher negative VOT values do not exhibit more regressive voicing assimilation than those showing lower VOT values.

3.2. *Interaction of phonetic properties*

Data on voicing degree and on segmental duration and intensity for the cluster pairs differing in the underlying C2 voicing status listed in Section 2.1(d) allow evaluating the relative power of the voicing phonetic properties.

3.2.1. *Voicing and segmental duration*

Statistical analysis results for the cluster pairs revealed differences in the temporal extent of vocal fold vibration during C2 and in C2 duration as a function of C2 underlying voicing. Indeed, underlyingly voiced consonants were found to exhibit not only a longer vocal fold vibration period but also a shorter duration than their voiceless counterparts for all three cluster types, i.e., stop + stop ($F(1,7)=40.87$, $p<0.001$, $F(1,7.1)=13.98$, $p<0.001$), stop + fricative ($F(1,7)=105.7$, $p<0.001$, $F(1,7)=104.00$, $p<0.001$) and fricative + stop ($F(1,7)=153.07$, $p<0.001$, $F(1,7)=18.27$, $p<0.004$). Differences in C2 duration as a function of underlying voicing are visible in Fig. 9 (black rectangles).

Moreover, the two phonetic characteristics were consistently transferred to C1. As shown in Fig. 10, all clusters exhibit significantly higher C1 voicing percentages before a voiced consonant (unfilled bars) than before a voiceless one (filled bars), i.e., stop + stop ($F(1,7.1)=68.8$, $p<0.001$), stop + fricative ($F(1,7)=154.7$, $p<0.001$), fricative + stop ($F(1,7)=33.90$, $p<0.001$). On the other hand, there was a significant regressive duration effect which was found not to proceed in the same way for all three cluster types. Regarding the stop + stop and fricative + stop sequences and as shown in Fig. 9, i.e., C1 (light gray rectangles) turned out to be significantly longer before a voiceless C2 than before a voiced C2 ($F(1,7.1)=21.75$, $p<0.01$, $F(1,7)=36.81$, $p<0.001$). Moreover, the fricative + stop sequences but not the stop + stop sequences exhibited an inverse relationship between the duration of the vowel preceding the cluster and that of C1 and C2 such that V (unfilled rectangles) was significantly longer in clusters with a voiced C2 than in those with a voiceless C2 in the case of the former cluster type vs the latter ($F(1,7)=57.07$, $p<0.001$). Also according to Fig. 9, a different duration pattern holds for stop + fricative clusters since both V and C1 turned out to be significantly longer before a voiced vs voiceless C2 in this case ($F(1,7)=66.20$, $p<0.001$; $F(1,7)=17.30$, $p<0.01$). A more robust inverse relationship pattern between C2 duration and V and C1 duration in stop + fricative clusters than in stop + stop and fricative + stop sequences may be related to the presence of larger differences in C2 duration as a function of underlying voicing for the former cluster type than for the two latter ones. Fig. 9 also reveals the presence of a greater overall VCC duration for clusters with a voiceless C2 than for those with a voiced C2.

The production mechanisms used by the individual speakers for the implementation of regressive voicing assimilation in

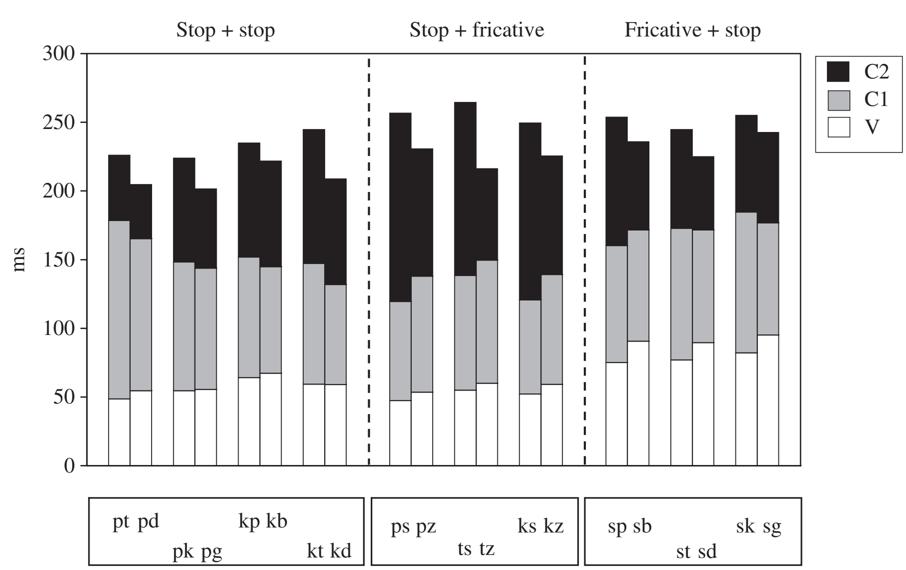

Fig. 9. V, C1 and C2 durations for cluster pairs differing in the C2 voicing status averaged across speakers.

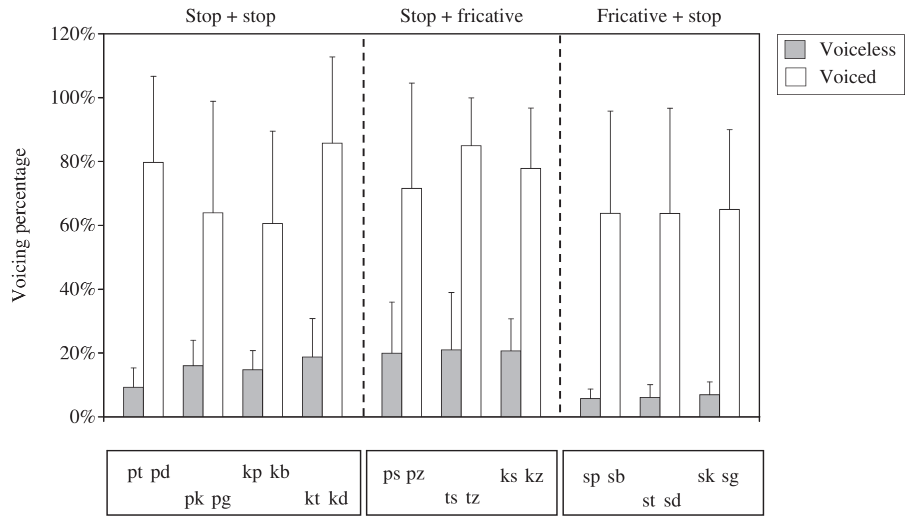

Fig. 10. C1 voicing percentages for cluster pairs differing in the C2 voicing status averaged across speakers. Error bars correspond to one standard deviation of the voicing means.

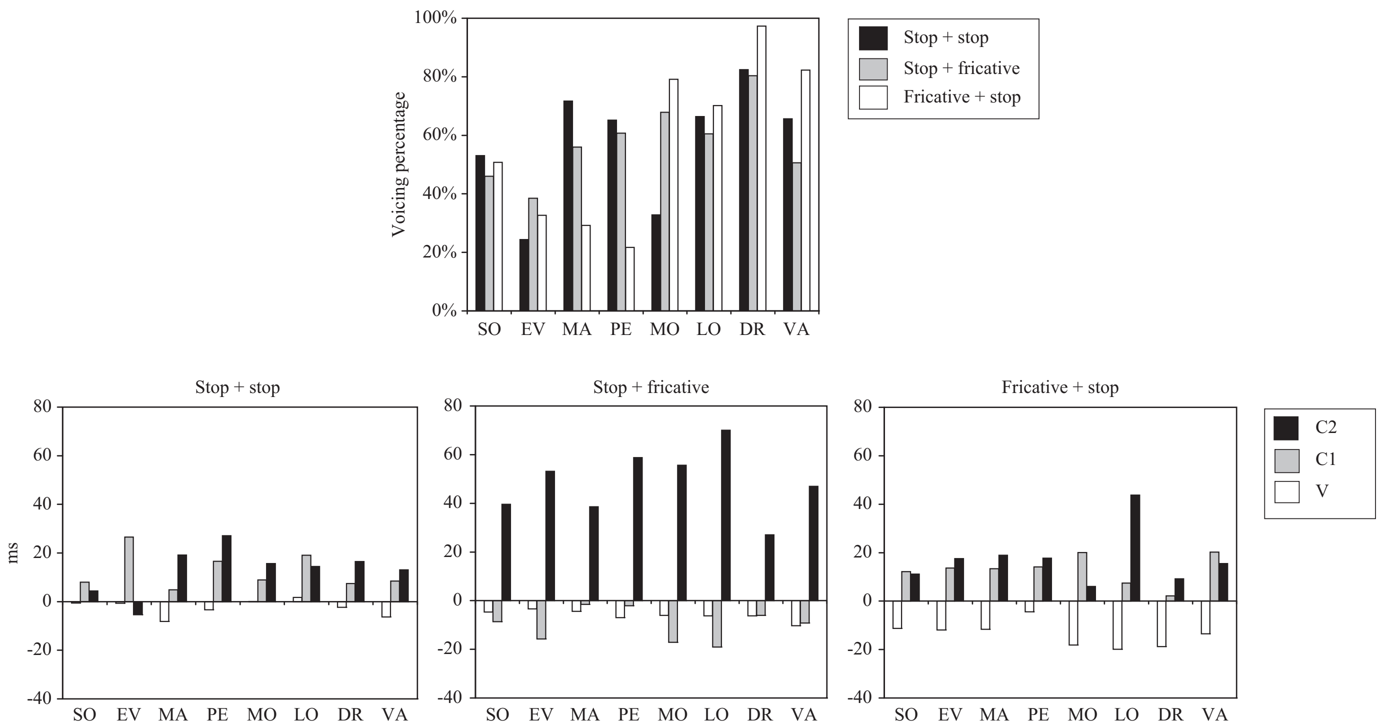

Fig. 11. (Top) Differences in C1 voicing degree as a function of the voiced minus voiceless C2 condition for all cluster types and speakers. (Bottom) Differences in V, C1 and C2 duration as a function of the voiceless minus voiced C2 condition for all cluster types and speakers.

clusters will be analyzed next. Fig. 11 (upper graph) displays differences in C1 voicing percentage between the voiced C2 and the voiceless C2 conditions, i.e., % before voiced C2 minus % before voiceless C2, plotted as a function of cluster type for all eight speakers. For all three cluster types and speakers, positive bars range between 20% and practically 100% and indicate that the vocal fold vibration period is clearly longer when C2 is underlyingly voiced than when it is underlyingly voiceless. Moreover, in parallel to the speaker-dependent data reported in Section 3.1.1, a significant C1 voicing × speaker interaction obtained in the ANOVAs run on the voicing data for the stop+stop condition ($F(4,47)=2.73$, $p<0.05$) and for the fricative+stop condition ($F(7,32)=22.39$, $p<0.001$) reveals that there is minimal voicing for some speakers (mostly EV, but also MA, PE and MO in the case of one cluster type) and maximal voicing for others (DR, VA).

While V, C1 and C2 happen to be significantly longer for some speakers than for others presumably due to differences in speech rate, ANOVAs performed on the cross-speaker segmental duration data yielded no significant interaction between the factor ‘speaker’ and either V, C1 or C2 duration, thus meaning that speakers do not differ among themselves regarding the effect of the underlying C2 voicing status on the duration of the preceding consonant and vowel segments. These statistical results are consistent with differences in V, C1 and C2 duration between the voiceless and voiced C2 conditions displayed in the three bottom graphs of Fig. 11 (in the graphs, the difference is positive if segments are longer in the voiceless vs voiced condition and negative if the opposite relationship holds). The graphs show the same scenario for all speakers: V (not C1) is generally longer before a voiced vs voiceless C2 in stop+stop and fricative+stop sequences, and the same contextual relationship applies to both V and C1 in stop+fricative clusters.

The finding that the factor 'speaker' interacts significantly with C1 voicing but not with V and C1 duration indicates that there is no speaker-dependent compensation between the temporal extent of vocal fold vibration and segmental duration. Thus, it is not the case that speakers with lower C1 voicing percentages exhibit larger V and C1 duration differences as a function of the C2 underlying voicing status. In summary, while degree of vocal fold vibration and segmental duration may be robust voicing phonetic properties, the latter parameter is kept more constant than the former one.

3.2.2. *Stop burst and frication noise characteristics*

In stop+stop sequences, neither the frequency of occurrence (64.2–93.4%) nor the duration (below 10 ms) of the C1 stop burst showed a main C2-dependent voicing effect or a significant voicing × speaker interaction.

Statistical tests run on the absolute and relative frication noise energy values in fricative+stop sequences yielded no main underlying voicing effect and a significant speaker × voicing interaction ($F(7,32)=4.65$, $p<0.001$; $F(7,32)=5.35$, $p<0.001$) according to which /s/ was more intense when occurring before a voiceless vs voiced consonant for some speakers, i.e., PE (3.03 dB difference), EV (2.95 dB), MO (2.38 dB), MA (2.02 dB) and SO (1.27 dB), but not for others, i.e., LO, DR and VA. Moreover, speaker-dependent differences in degree of vocal fold vibration and energy level between clusters with a voiced and a voiceless C2 turned out to be negatively correlated ($r=-0.67$), thus meaning that speakers may compensate for the lack of voicing during /s/ before a voiced C2 by increasing the energy level of the frication noise in the voiceless C2 condition. This compensatory behavior is shown in Fig. 12. In the figure, normalized percentage differences in voicing (filled bars) and energy level (unfilled bars) reveal that there may be a compensatory relationship between the two measures for some speakers, i.e., minimal voicing – maximal noise energy for EV, MA and PE and maximal voicing – minimal noise energy for LO, DR and VA.

4. Summary and discussion

In view of the large duration range of the vocal fold vibration period in the case of clusters with an obstruent C1 and a voiced C2, data presented in Section 3 indicate that regressive voicing adaptation may be far from complete in Catalan. Moreover, C1 voicing percentages appear to be less than in other Romance languages, i.e., voicing during /p, t, k/ before a voiced C2 amounts to 58.7% in Catalan and to 76.3% in French (Snoeren et al., 2006). Two major findings seeking to account for gradience in C1 obstruent voicing have been reported which will be discussed later in this section:

(a) consonants proceed according to the coarticulatory resistance scale predicted by the DAC model as a general rule, and there is a close relationship between coarticulatory resistance and coarticulatory aggressiveness for several consonants;
(b) data for cluster pairs differing in the underlying C2 voicing status reveal that regressive voicing may be considered an assimilatory process in Catalan which is signaled not only by vocal fold vibration during C1 but also by C1 duration and intensity and by preceding vowel duration.

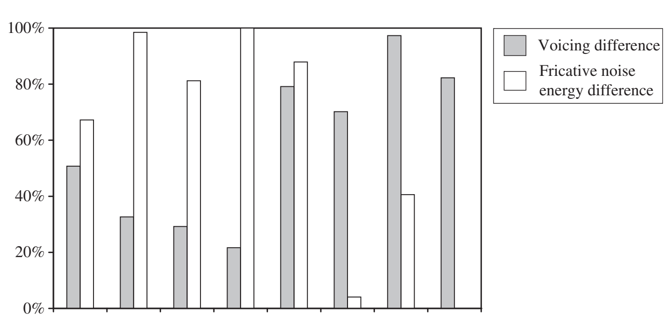

Fig. 12. Differences in /s/ voicing degree and frication noise energy as a function of the C2 voicing status in fricative+stop sequences for all speakers. See text for details.

In agreement with our initial hypothesis, voicing coarticulation resistance turned out to be higher for nasals, laterals and approximants, including the allophones of /b, d, g/, than for stops, fricatives and the alveolar trill. An explanation for these consonant-dependent differences in coarticulatory resistance may be sought in the fact that voicing is facilitated by an unimpeded continuous airflow for sonorants, and becomes harder to maintain due to an intraoral pressure rise for obstruents and the trill and by conflicting requirements between vocal fold approximation and enough glottal opening for airflow for fricatives. Among obstruents, the fricative /z/ exhibited as much or more voicing than the underlying voiced stops at the C2 site, while fricatives turned out to voice less than stops at the C1 site. In line with data for other languages (see Section 1), this position-dependent difference in voicing behavior may be attributed to a trend for fricatives to allow for more devoicing than stops unless positively specified for voicing. Also as predicted, back stop articulations were found to be more prone to show periods of voicelessness during their production than more anterior ones due to differences in back cavity size and vocal tract compliance. The labiodental fricative was reported to exhibit a special voicing behavior: voicing percentages were lower for /f/ than for /s, ʃ/ in intervocalic and in C1 position in clusters in spite of /f/ being more anterior and exhibiting a lower noise intensity level than /s, ʃ/. A possible explanation for this voicing difference may be sought in the need to enhance the labiodental fricative perceptually (see Section 1.1.2.1). In support of this possibility and contrary to reports for other languages, intervocalic word final /f/ (92.9 ms, sd=20.5) happened to be especially long when compared to the other fricatives /s/ (70.8 ms, sd=15.1) and /ʃ/ (88.6 ms, sd=20.5), and there was a trend for Catalan speakers with a longer /f/ to exhibit less voicing than those with shorter realizations of the consonant (a correlation analysis between duration and voicing for intervocalic word final /f/ yielded an r value of -0.68). In fact, current descriptions assign a voiceless realization to intervocalic word final /f/ in Catalan oxytones while lingual fricatives become voiced word finally before a word initial vowel (Recasens, 1993; see Section 1.1.2.1). One reviewer has pointed out to us that the failure for /f/ to exhibit much voicing may be attributed to the absence of /v/ in the phoneme inventory of Catalan and, more generally, to a structure preservation principle which prevents phonological rules from applying if they generate structures which are prohibited underlyingly (Kiparsky, 1985; Steriade, 1995). This possibility does not sound too feasible, however, since intervocalic word final /f/ also fails to undergo voicing in Catalan dialects (e.g., Majorcan Catalan) where /v/ has phonemic status, and structure preservation is supposed to apply lexically while the rule of interest applies postlexically in Catalan. We believe instead that the perceptual enhancement of /f/ may be partly due to the fact that there are only about four genuine and widely used Catalan words showing this consonant in postvocalic word final position (a very different scenario from that for other languages like English).

Results on voicing adaptation reported in the present study indicate a close relationship between the degree of voicing in the triggering and the target consonants (and thus, between coarticulatory resistance and coarticulatory aggressiveness) in a good number of instances. This relationship accounts for more regressive voicing during C1 before /z/ than before voiced stops and /r/, and before approximant vs stop allophones of /b, d, ɡ/, and for more C2 devoicing after fricatives vs stops. Regressive and progressive effects in vocal fold vibration for stops and fricatives are reminiscent of findings for other languages (see Section 1), and may be interpreted based on the voicing requirements and the amplitude of the voiceless gesture for the two sets of consonants: differences in regressive adaptation appear to follow from the fact that active voicing imposes lesser aerodynamic constraints on C1 when C2 is a voiced fricative than when it is a voiced stop; on the other hand, voiceless fricatives trigger more progressive adaptation than unaspirated voiceless stops since the amplitude of the devoicing gesture is greater for the former vs the latter.

There were cases where no positive relationship between voicing in the triggering and the target consonant is envisaged and which therefore run against the assumptions of the DAC model: no coarticulatory voicing effects associated with place of articulation for C2 = /b, d, ɡ/ (regressive) and for C1 = /p, t, k/ and /f, s, ʃ/ (progressive) were observed; there was little regressive voicing in clusters where stops and fricatives were followed by nasals and laterals, which exhibit maximal voicing in the C2 position. The latter finding deserves closer attention. According to current explanations advocated by phonologists (see Section 1.1.2.2), sonorants are specified for voicing albeit less actively than voiced obstruents, the former involving just a particular glottal state induced spontaneously by a low supraglottal pressure and the latter several articulatory adjustements. In our view, the specific voicing behavior of these clusters is to be sought, at least in part, in the aerodynamic and articulatory demands involved: the absence of much regressive voicing appears to be required in order to avoid the presence of nasal airflow for a nasal C2 and the anticipatory oral constriction gesture for a lateral C2 during the preceding obstruent, which would impair the high intensity frication noise for fricatives and hinder the intraoral pressure build-up for the generation of a prominent burst for stops. Our explanation is consistent with two findings reported in our study: other sonorants, i.e., the approximants /j/ and [β, ð, ɣ], may cause a high degree of C1 voicing to occur; C1 shortens to a larger extent before nasals than before other consonants, i.e., mean C1 duration was about 70–80 ms before /m,n/ and 80–95 ms before /b, d, ɡ, l, z, r, ʎ, j/, which in the case of a fricative C1 may be attributed to encroachment of velar opening and drop of noise source pressure (Haggard, 1973; Solé, 2007). In any case, a careful phonetic analysis of the production conditions operating in languages where obstruents but not sonorants trigger regressive voicing assimilation onto preceding obstruents (e.g., Russian, German, Dutch; see Section 1.1.2.2) needs to be carried out in order to verify the universal validity of the DAC model predictions regarding consonant voicing adaptation in clusters.

Catalan speakers were found to agree in several ways regarding the extent to which their voicing adaptation data conform to the predictions of the DAC model. Most speakers behaved as expected by showing more regressive voicing as a function of /z/ than of stops and the trill, and more progressive C2 devoicing after fricatives than after stops. They also was also agreement among most speakers in exhibiting less regressive voicing than expected for nasals and laterals.

Three other findings regarding the degree of vocal fold vibration for consonants in clusters need to be pointed out. Firstly, the fact that progressive devoicing may occur in a language with voiced stops with voicing lead like Catalan. The fact that the degree to which C2 may devoice (Fig. 7, right graph) is generally less than the C1 voicing percentages (Fig. 5, left graph) reveals that progressive devoicing effects are less prominent than regressive voicing effects and thus merely coarticulatory, which is consistent with the regressive nature of the voicing assimilation process in Catalan. Secondly, the degree of vocal fold vibration during C1 varies with speaker, with some speakers exhibiting little voicing and others considerable voicing. It cannot be discarded that speaker-dependent differences in voicing degree may be dialect-dependent with subjects from Western Catalan and urban Barcelona favoring the voicing of obstruents less than speakers born in smaller towns from the Eastern Catalan dialect domain (see Section 1). Thirdly, speaker-dependent differences in vocal fold vibration during C1 appear to proceed independently of the corresponding negative VOT values for voiced stops. This finding suggests that, while Westbury's hypothesis regarding the relation between voicing in single consonants and voicing adaptation in clusters may be at work in languages with vs without voicing lead, it does not seem to hold when data for the individual speakers of a given language are taken into consideration.

A discussion about the interactive contribution of several voicing phonetic characteristics follows. Obstruent + obstruent cluster pairs show significant differences as a function of the C2 voicing status, not only in C1 voicing degree but also in V and/or C1 duration. In particular, C1 turned out to be longer before a voiceless C2 than before a voiced C2 in stop + stop and fricative + stop sequences, V was longer in clusters with a voiced vs voiceless C2 in fricative + stop clusters, and both V and C1 turned out to be significantly longer when C2 was voiced than when it was voiceless in stop + fricative sequences. The finding that the segmental duration patterns just referred to remain more constant than vocal fold vibration across cluster types and speakers suggests that the former characteristic may be a more robust voicing cue than the latter and hence under active control (as for the relationship between the lack of variability across conditions for a given phonetic characteristic and the speaker's control of articulatory events, see Solé and Ohala, 2010 and Solé, 1995). However, since in spite of these variability differences the two parameters, i.e., vocal fold vibration and segmental duration, turned out to be significant for all cluster types, we prefer to conclude that they are both effective in signaling the voiced/voiceless distinction. In support of this possibility, no compensatory trend between the period of vocal fold vibration and segmental duration was found to hold for the speakers under study, i.e., it was not the case that the cueing power of segmental duration increased whenever vocal fold vibration became less prominent. This finding is not entirely in agreement with the traditional view that vocal fold vibration should be the primary voicing cue in a language exhibiting voiced stops with voicing lead like Catalan nor with the proposal that stop closure and preceding vowel duration should prevail upon vocal fold vibration in cueing voicing (see Section 1.2). Instead, it appears that both phonetic properties, i.e., voicing fold vibration and segmental duration, may play a relevant role in marking the C1 voicing distinction in clusters undergoing regressive voicing assimilation. Perception data are needed in order to verify this possibility.

This article reports other relevant findings regarding the interactive role of several acoustic characteristics in the implementation of regressive voicing in Catalan clusters. Firstly, there are signs that speakers may compensate for the lack of voicing with the energy level of the frication noise in /sC/ sequences, i.e., /s/ intensity differences between clusters with a voiced and a

Table A1
List of consonant clusters and sentences with Catalan orthographic representation and English translation.

(a) Obstruent+voiced C2 clusters		
C1=/p/		
1. /pd/	no em donis el xarop dolç	"do not give me the sweet syrup"
2. /pɡ/	de la botiga en rep gots	"he/she receives glasses from the shop"
3. /pm/	he caçat un catxap maco	"I have hunted a nice young rabbit"
4. /pn/	han caçat un catxap nou	"they have hunted a new young rabbit"
5. /pl/	de la granja ell en rep làctics	"he gets dairy products from the farm"
6. /pz/	amb l'ullera percep zebres	"he/she views zebras with a telescope"
7. /pr/	ja no em queda cap ram	"I have no branch left"
8. /pʎ/	l'ampolla té un tap llarg	"the bottle has a long cork"
9. /pj/	de regal sempre rep iots	"he/she always receives yachts as a gift"
C1=/t/		
10. /tb/	es tracta d'un debat bàsic	"it is a basic debate"
11. /tɡ/	es presentà un soldat guerxo	"a one-eyed soldier showed up"
12. /tm/	m'he posat un calçat maco	"I am wearing nice footwear"
13. /tn/	m'he posat un calçat nou	"I am wearing new footwear"
14. /tl/	es tracta d'un soldat laic	"he happens to be a secular soldier"
15. /tz/	un valent soldat zulú	"a brave Zulu soldier"
16. /tr/	vaig servir de soldat ras	"I served as a common soldier"
17. /tʎ/	endevina aquest mot llarg	"guess this long word"
18. /tj/	es va comprar aquells set iots	"he/she bought those seven yatchs for himself/herself"
C1=/k/		
19. /kb/	això sí que és tabac bo	"this is indeed good tobacco"
20. /kd/	ell li ha donat un toc dolç	"he has given it a sweet buzz"
21. /km/	he comprat un tabac maco	"I have bought nice tobacco"
22. /kn/	he comprat un tabac nou	"I have bought new tobacco"
23. /kl/	d'aquesta ampolla en trec laca	"I take lacquer out of this bottle"
24. /kz/	jo del parc ara en trec zebres	"I take zebras out of the zoo"
25. /kr/	es tracta d'un suec ros	"he happens to be a blond Swede"
26. /kʎ/	partiren aquest soc llarg	"they choped this long trunk"
27. /kj/	aquest és un tabac ianqui	"this is American tobacco"
C1=/f/		
28. /fb/	a la cuina hi ha un xef basc	"there is a Basque chef at the kitchen"
29. /fd/	d'allí en sortia un baf dolç	"a sweet steam was coming from there"
30. /fɡ/	a la cuina hi ha un xef gal	"there is a French chef at the kitchen"
31. /fm/	a la cuina hi ha un xef maco	"there is a nice chef at the kitchen"
32. /fn/	preferiria un xef nou	"I would prefer a new chef"
33. /fl/	contractàrem un xef laic	"we hired a secular chef"
34. /fz/	prova els plats del xef zulú	"taste the Zulu chef's dishes"
35. /fr/	a la cuina hi ha un xef ros	"there is a blond chef at the kitchen"
36. /fʎ/	a la cuina hi ha un xef llarg	"there is a tall chef at the kitchen"
37. /fj/	contractàrem un xef ianqui	"we hired an American chef"
C1=/s/		
38. /sb/	aquell fou un envàs bo	"it was a good container"
39. /sd/	el fiscal tingué un cas dolç	"the public prosecutor had an easy case"
40. /sɡ/	ells m'han venut un gos guerxo	"they have sold me a one-eyed dog"
41. /sm/	es tracta d'un envàs moll	"it is a wet container"
42. /sn/	han patentat l'envàs nou	"they have patented the new container"
43. /sl/	això ha estat un fracàs lògic	"it has been an obvious failure"
44. /sr/	per aquí hi passa el gas ras	"the gas passes over the soil surface"
45. /sʎ/	l'atraparà amb el pas llarg	"he/she will catch him/her by walking fast"
46. /sj/	Coca Cola en envàs ianqui	"Cola Cola in an American container"
C1=/ʃ/		
47. /ʃb/	no hi arribo al calaix baix	"I cannot reach the low drawer"
48. /ʃd/	canvia'm aquest peix dolç	"give me another sweet fish"
49. /ʃɡ/	a la cuina esbandeix gots	"he/she rinses glasses at the kitchen"
50. /ʃm/	no l'agafis el feix moll	"do not take the wet bundle"
51. /ʃn/	agafa'l aquest feix nou	"take this new bundle"
52. /ʃl/	aquest sí que és un peix lent	"this is a slow fish indeed"
53. /ʃr/	no hi arribo al calaix ras	"I cannot reach the low drawer"
54. /ʃʎ/	necessitem un feix llarg	"we need a long bundle"
55. /ʃj/	el meu germà coneix ianquis	"my brother knows American people"
(b) VOT		
56. /p/	passa un soldat curd	"a Kurdish soldier goes by"
57. /b/	balla amb el xef curd	"she dances with the Kurdish chef"
58. /t/	talla aquell peix car	"slice that expensive fish"
59. /d/	dalt hi ha un sac tort	"there is a twisted sack upstairs"
60. /k/	caça un catxap curt	"he/she hunts a short young rabbit"
61. /ɡ/	gasten l'envàs car	"they consume the expensive container"
(c) Intervocalic fricatives		
62. /f/	hi treballava un xef alt	"a tall chef was working there"
63. /s/	camina amb el pas ample	"walk fast"
64. /ʃ/	omple el calaix ample	"fill the wide drawer"

Table A1 (*continued*)

(d) Cluster pairs		
Stop+stop		
65. /pt/	no queda cap talp	"there isn't any mole left"
66. /pd/	no queda cap dau	"there isn't any dice left"
67. /pk/	no queda cap cos	"there isn't any body left"
68. /pg/	no queda cap got	"there isn't any glass left"
69. /kp/	té un toc pobre	"it has a poor touch"
70. /kb/	té un toc bosni	"it has a Bosnian touch"
71. /kt/	ha quedat poc tou	"it has come out not smooth enough"
72. /kd/	ha quedat poc dolç	"it has come out not sweet enough"
Stop+fricative		
73. /ps/	no hi ha cap ceba	"there aren't any onions"
74. /pz/	no hi ha cap zebra	"there aren't any zebras"
75. /ts/	l'ha tret tot sec	"he/she has taken it out completely dry"
76. /tz/	ha tret tot zeros	"he has got zero grades only"
77. /ks/	hi posa poc seny	"he/she is not conscientious enough"
78. /kz/	hi posa poc zel	"he/she is not zelous enough"
Fricative+stop		
79. /sp/	fou un cas pobre	"it was a poor case"
80. /sb/	fou un cas bosni	"it was a Bosnian case"
81. /st/	porta un pas tort	"he/she walks in a twisted manner"
82. /sd/	porta un pas dolç	"he/she walks gently"
83. /sk/	fou un cas car	"it was an expensive case"
84. /sg/	fou un cas gal	"it was a French case"

voiceless C2 were largest for those speakers who showed least vocal fold vibration during /s/ before a voiced consonant. Therefore, it seems that speakers may signal the voicing distinction by increasing the /s/ frication noise intensity level whenever the vocal fold vibration period does not play a sufficient relevant role. Secondly, a comparison with other languages indicates that differences in vowel duration as a function of C2 voicing are similar to those found in French and shorter than those occurring in English (Chen, 1970; Snoeren et al., 2006), i.e., in Catalan, the shorter vowel is only 15% and 10% less than the longer counterpart in fricative+stop and stop+fricative clusters, respectively. Moreover, the degree to which V and/or C1 duration depends on C2 voicing is conditioned by consonant manner of articulation, i.e., segmental duration contributes mostly to marking C1 voicing when the cluster contains a fricative in fricative+stop and stop+fricative sequences than when it does not (see Section 1.2 regarding the effect of fricative vs stop manner of articulation on preceding vowel duration in other languages).

To summarize, the initial predictions of the DAC model regarding voicing adaptation in heterosyllabic two-consonant clusters were found to apply in the case of several consonant sequences and a subset of speakers. The model predicts that there ought to be a positive relationship between voicing degree in a given consonant and the extent to which this consonant causes an adjacent consonant to acquire voicing. In comparison to the lingual coarticulation scenario, a problem with extending the DAC model to vocing coarticulation appears to lie on devoicing effects which may result from specific interarticulatory maneuvers such as those occurring in fricative+nasal clusters. As reported for other Romance languages elsewhere, regressive voicing adaptation effects were found to exceed progressive voicing effects which is in agreement with the notion that voicing assimilation is a regressive process in Catalan. Contrary to the formulation of the phonological rule, however, regressive voicing was far from complete when vocal fold vibration data were taken into consideration. In partial agreement with languages where voiced stops are not implemented through considerable voicing lead, speakers of Catalan appear to rely on other phonetic characteristics besides vocal fold vibration such as segmental duration and intensity for signaling the C1 voicing distinction. This pattern of phonetic behavior is also consistent with the lack of correspondence between negative VOT for voiced stops and degree of vocal fold vibration during C1 for the individual speakers' data subjected to analysis in the present investigation.

Acknowledgments

This research was supported by Project FFI2009-09339 of the Spanish Ministry of Science and Innovation and FEDER, by the ICREA Academia program, and by Project 2009SGR3 of the Catalan government. We would like to thank three anonymous reviewers and the Associate Editor Taehong Cho for comments, Alexis Michaud for allowing us to use Peakdet, and Sergi Abadal for writing the Peakdet 2 version of the MatLab script.

Appendix

See Table A1.

References

Abadal, S., & Recasens, D. (2009). Peakdet2. An instructions manual 〈http://voiceresearch.free.fr/egg〉.

Abdelli-Beruh, N. B. (2004). The stop voicing contrast in French sentences: Contextual sensitivity of vowel duration, closure duration, voice onset time, stop release and closure voicing. *Phonetica*, *61*, 201–219.

Balise, R. R., & Diehl, R. L. (1994). Some distributional facts about fricatives and a perceptual explanation. *Phonetica*, *51*, 99–110.

Beckman, J., Jessen, M., & Ringen, C. (2009). German fricatives: Coda devoicing or positional faithfulness? *Phonology*, *26*, 231–268.

Bonet, E., & Lloret, M. R. (1998). *Fonologia catalana*. Barcelona: Ariel.

Burton, M. W., & Robblee, K. E. (1997). A phonetic analysis of voicing assimilation in Russian. *Journal of Phonetics*, *25*, 97–114.

Chen, M. (1970). Vowel length variation as a function of the voicing of the consonant environment. *Phonetica*, *22*, 129–159.

Cho, T., Jun, S.-A., & Ladefoged, P (2002). Acoustic and aerodynamic correlates of Korean stops and fricatives. *Journal of Phonetics*, *30*, 193–228.

Crystal, T. H., & House, A. S. (1988). The duration of American-English stop consonants: An overview. *Journal of Phonetics*, *16*, 285–294.

Cuartero, N. (2001). *Voicing assimilation in Catalan and English*. Ph.D. Dissertation, Universitat Autònoma de Barcelona.

Docherty, G. J. (1992). *The timing of voicing in British English obstruents*. Berlin-New York: Foris.

Dommelen, W. A. van (1983). Some observations on assimilation of voicing in German and Dutch. In: M. van den Broecke, V. van Heuven, & W. Zonneveld (Eds.), *Studies for Antonie Cohen. Sound structures* (pp. 47–56). Dordrecht: Foris Publications.

Dorman, M., Studdert-Kennedy, M., & Raphael, L. (1977). Stop consonant recognition: Release bursts and formant transitions are functionally equivalent, context-sensitive cues. *Perception and Psychophysics*, *22*, 109–122.

Dvorak, V. (2010). Voicing assimilation in Czech. *Rutgers Working Papers in Linguistics*, *3*, 115–144.

Haggard, M. P. (1973). Abbreviation of consonants in English pre- and post-vocalic clusters. *Journal of Phonetics*, *1*, 9–24.

Haggard, M. P. (1978). The devoicing of voiced fricatives. *Journal of Phonetics*, *6*, 95–102.

Hall, D. C. (2007). *The role and representation of contrast in phonological theory*, Ph.D. Thesis, Department of Linguistics, University of Toronto.

Hallé, P. A., & Adda-Decker, M. (2007). Voicing assimilation in journalistic speech. In: J. Trouvain, & W. J. Barry (Eds.), *16th International congress of phonetic sciences* (pp. 493–496). Saarbrücken: Universität des Saarlandes.

Henrich, N., d'Alessandro, C., Doval, B., & Castellengo, M. (2004). On the use of the derivative of electroglottographic signals for characterization of nonpathological phonation. *Journal of the Acoustical Society of America*, *115*, 1321–1332.

Hoole, P. (1999). Laryngeal coarticulation. In: W. J. Hardcastle, & N. Hewlett (Eds.), *Coarticulation: Theory, data and techniques* (pp. 105–121). Cambridge: Cambridge University Press.

Jansen, W. (2004). *Laryngeal contrast and phonetic voicing: A laboratory phonology approach to English, Hungarian, and Dutch*. Ph.D. Dissertation, University of Groningen.

Jessen, M. (2004). Instability in the production and perception of intervocalic closure voicing as a cue to *bdg* vs *ptk* in German. *Folia Linguistica*, *38*, 27–41.

Jesus, L. M., & Shadle, C. H. (2002). Analysis of Portuguese fricatives. *Journal of Phonetics*, *30*, 437–464.

Kiparsky, P. (1985). Some consequences of Lexical Phonology. *Phonology Yearbook*, *2*, 82–138.

Kohler, K. (1979). Parameters in the production and the perception of plosives in German and French. *Arbeitsberichte des Institut für Phonetik uns Digitale Sprachverarbeitung, Kiel University*, *12*, 261–280.

Kohler, K. (1984). Phonetic features in phonology: The feature fortis-lenis. *Phonetica*, *41*, 150–174.

Kulikov, V. (2011). *The phonetics and phonology of voice assimilation and sonorant transparency in normal and fast speech in Russian*. Manuscript, University of Iowa.

Laeufer, C. (1992). Patterns of voicing-conditioned vowel duration in French and English. *Journal of Phonetics*, *20*, 411–440.

Lisker, L., & Abramson, A. S. (1964). A cross-language study of voicing in initial stops: Acoustical measurements. *Word*, *20*, 384–422.

Löfqvist, A., & McGarr, N. (1987). Laryngeal dynamics in voiceless consonant production. In: T. Baer, C. Sasaki, & K. S. Harris (Eds.), *Laryngeal function in phonation and respiration* (pp. 391–402). San Diego: College Hill.

Lulich, S. (2004). Russian [v]: An acoustic study. *Folia Linguistica*, *28*, 63–81.

Markó, A., Gráczi, T. E., & Bóna, J. (2010). The realisation of voicing assimilation rules in Hungarian spontaneous and read speech: Case studies. *Acta Linguistica Hungarica*, *57*, 210–238.

Martínez Celdrán, E., & Fernández Planas, A. M. (2007). *Manual de fonética española*. Barcelona: Ariel.

Mascaró, J. (1987). Underlying voicing recoverability of finally devoiced obstruents in Catalan. *Journal of Phonetics*, *15*, 183–186.

Mascaró, J., & Rafel, J. (1990). *Diccionari català invers amb informació morfològica*. Barcelona: Publicacions de l'Abadia de Montserrat.

Mazaudon, M., & Michaud, A. (2008). Tonal contrasts and initial consonants: A case study of Tamang, a 'missing link' in tonogenesis. *Phonetica*, *65*, 231–256.

Myers, S. (2002). *Gaps in factorial typology: The case of voicing in consonant clusters*. Manuscript, University of Texas at Austin.

Ohala, J. J., & Riordan, C. J. (1979). Passive vocal tract enlargement during voiced stops. In: J. J. Wolf, & D. Klatt (Eds.), *Speech communication papers* (pp. 89–92). New York: Acoustical Society of America.

Ohala, J. J., & Solé, M. J. (2010). Turbulence and phonology. In: S. Fuchs, M. Toda, & M. Zygis (Eds.), *Turbulent sounds. An interdisciplinary guide* (pp. 37–97). Berlin: Mouton de Gruyter.

Pinho, C. M. R., Jesus L. M. T., & Barney, A. (2009). Aerodynamics of fricative production in European Portuguese. In *Proceedings of the Interspeech 2009*, pp. 472–475.

Pirello, K. J., Blumstein, S. E., & Kurowski, K. (1997). The characteristics of voicing in syllable–initial fricatives in American English. *Journal of the Acoustical Society of America*, *101*, 3754–3765.

Recasens, D. (1986). *Estudis de fonética experimental del català oriental central*. Barcelona: Publicacions de l'Abadia de Montserrat.

Recasens, D. (1993). *Fonètica i fonologia*. Barcelona: Enciclopèdia Catalana.

Recasens, D., & Espinosa, A. (2006). Estudi experimental de les consonants fricatives del mallorquí i del valencià. *Estudis Romànics*, *28*, 125–150.

Recasens, D., & Espinosa, A. (2007). Phonetic typology and positional allophones for alveolar rhotics in Catalan. *Phonetica*, *63*, 1–28.

Recasens, D., & Espinosa, A. (2009). An articulatory investigation of lingual coarticulatory resistance and aggressiveness for consonants and vowels in Catalan. *Journal of the Acoustical Society of America*, *125*, 2288–2298.

Recasens, D., Pallarès, M. D., & Fontdevila, J. (1997). A model of coarticulation based on articulatory constraints. *Journal of the Acoustical Society of America*, *102*, 544–561.

Rice, K. (1993). A reexamination of the feature [sonorant]: The status of 'sonorant obstruents'. *Language*, *69*, 308–344.

Rothenberg, M., & Mahshie, J. J. (1988). Monitoring vocal fold abduction through vocal fold contact area. *Journal of Speech and Haring Research*, *312*, 338–351.

Slis, I. H. (1981). Rules for assimilation of voice in Dutch. In: R. Channon, & L. Shockey (Eds.), *In honor of Ilse Lehiste* (pp. 225–239). Dordrecht: Foris Publications.

Slis, I. H. (1986). Assimilation of voice in Dutch as a function of stress, word boundaries and sex of speaker and listener. *Journal of Phonetics*, *14*, 311–326.

Smith, C. L. (1997). The devoicing of /z/ in American English: Effects of local and prosodic context. *Journal of Phonetics*, *25*, 471–500.

Snoeren, N. D., Hallé, P. A., & Seguí, J. (2006). A voice for the voiceless: Production and perception of assimilated stops in French. *Journal of Phonetics*, *34*, 241–268.

Solé, M. J. (1995). Spatio-temporal patterns of velo-pharyngeal action in phonetic and phonological nasalization. *Language and Speech*, *38*, 1–23.

Solé, M. J. (2007). The stability of phonological features within and across segments: The effect of nasalization on frication. In: P. Prieto, J. Mascaró, & M. J. Solé (Eds.), *Segmental and prosodic issues in Romance phonology* (pp. 41–66). Amsterdam: John Benjamins.

Solé, M. J., & Ohala, J. J. (2010). What is and what is not under the control of the speaker. Intrinsic vowel duration. In: C. Fougeron, B. Kühnert, M. Imperio, & N. Vallée (Eds.), *Laboratory phonology*, 10 (pp. 607–655). Berlin: de Gruyter.

Steriade, D. (1995). Underspecification and markedness. In: Goldsmith (Ed.), *Handbook of phonological theory* (pp. 114–174). Oxford: Blackwell.

Stevens, K. N., Blumstein, S., Glicksman, S., Burton, M., & Kurowski, K. (1992). Acoustical and perceptual characteristics of voicing in fricatives and fricative clusters. *Journal of the Acoustical Society of America*, *91*, 2979–3000.

Westbury, J. (1975). The status of regressive voicing assimilation as a rule of Russian. *Texas Linguistic Forum*, *1*, 131–144.

Westbury, J. (1979). Aspects of the temporal control of voicing in consonant clusters in English. *Texas Linguistic Forum*, *14* Austin, Texas: Department of Linguistics, University of Texas at Austin.

Westbury, J., & Keating, P. (1986). On the naturalness of consonant voicing. *Journal of Linguistics*, *22*, 145–166.

Zue, V. (1980). *Acoustic characteristics of stop consonants. A controlled study*. Bloomington: Indiana Linguistic Club.