Original Paper

Phonetica 2014;71:128–156
DOI: 10.1159/000369028

Received: June 5, 2014
Accepted after revision: October 7, 2014
Published online: November 17, 2014

Gradient Phonetic Implementation of Regressive Voicing Assimilation in Catalan Heterosyllabic Two- and Three-Consonant Clusters

Daniel Recasens

Universitat Autònoma de Barcelona and Institut d'Estudis, Catalans, Barcelona, Spain

Abstract

Electroglottographic data for word-final obstruents in C#C and CC#C sequences with a word-initial voiced consonant indicate that regressive voicing adaptation is a categorical and thus assimilatory process for most Catalan speakers. Word-final obstruents are planned as voiced since they exhibit full voicing or else an initial voicing period which is longer than the voicing lag associated with the vowel preceding the cluster. Segmental duration may also be used by speakers for realizing word-final obstruents as voiced in C#C but not in CC#C clusters. The phonetic implementation of voicing assimilation proceeds gradually: voicing for word-final obstruents differs considerably among speakers and is less for fricatives than for stops and whenever the word-initial consonant is an obstruent or an approximant than when it is a nasal, a lateral or an alveolar trill. The study also reveals that those C1 realizations which are less prone to acquire voicing show more token-to-token variability in voicing and segmental duration and therefore appear to be less tightly controlled by speakers. In conjunction with data from other studies, these Catalan data suggest that speakers and languages with prevoiced stops may differ with respect to the more or less gradient phonetic implementation of the regressive voicing assimilation process in heterosyllabic consonant clusters.

1 Introduction

Catalan, the language subjected to investigation in the present study as well as other languages with voiced stops with negative voice onset time (also referred to commonly as languages with stop prevoicing) such as French, Hungarian and Slovak are said to share a regressive voicing assimilation rule according to which word-final obstruents become phonetically voiced when followed by all or a subset of voiced consonants in heterosyllabic C#C sequences (see Wheeler, 2005, for Catalan). This voicing

KARGER

E-Mail karger@karger.com
www.karger.com/pho

0031–8388/14/0712–0128
$39.50/0

Daniel Recasens
Departament de Filologia Catalana
Universitat Autònoma de Barcelona
Bellaterra, ES–08193 Barcelona (Spain)
E-Mail daniel.recasens@uab.es

adaptation process is in contrast with that occurring in English and other languages where voiced stops do not regularly show negative voice onset time and heterosyllabic C#C sequences may exhibit a mixed assimilation scenario involving regressive voicing, progressive voicing or no voicing adaptation between the two consonants (Westbury, 1979; Docherty, 1992). It should be added that languages with regressive voicing assimilation may differ as to whether they neutralize or maintain the voicing contrast between underlyingly voiced and voiceless obstruents word-finally: while French and Hungarian do not neutralize the contrast, Catalan, Russian, Czech, Polish, Slovak and Western Flemish do and therefore exhibit final devoicing as well (Wetzels and Mascaró, 2001).

The goal of the present investigation is to use electroglottographic (EGG) data on heterosyllabic C#C and CC#C clusters from previous publications (Recasens and Mira, 2012, 2013) in order to explore more in depth the phonetic validity of the Catalan regressive voicing assimilation process. In Catalan C#C sequences, stops and fricatives are supposed to assimilate in voicing to any following voiced consonant (/tz/, [dz] *se**t** **z**eros* 'seven zeros'; /sm/, [zm] *go**s** **m**oll* 'wet dog'). CC#C sequences, on the other hand, may be split according to whether C2 is a fricative or a stop: in CC#C sequences with a fricative C2, the fricative is expected to be realized as voiced and so C1 if it is a stop (/msb/[1], [mzβ] *ra**ms** **b**ons* 'good branches'; /psb/, [bzβ] *ca**ps** **b**ons* 'good heads'); in CC#C sequences with a stop C2, the stop ought to be realized as voiced and so C1 if it is a fricative (/rpd/, [rbd] *se**rp** **d**aurada* 'golden snake'; /skd/, [zgd] *ca**sc** **d**ur* 'tough helmet'). In consonant sequences where the word-initial consonant is voiceless underlyingly, the word-final obstruents are also phonetically voiceless (/gp/, [kp] *un cé**c** **p**obre* 'a poor blind man'; /gsk/, [ksk] *són cé**cs** **c**ultes* 'they are educated blind men').

The research issues of the present study differ from those of our two previous publications on voicing assimilation in Catalan C#C and CC#C sequences with a word-initial voiced consonant in important respects. Recasens and Mira (2012, 2013) investigate the degree of voicing adaptation using cross-token voicing ratios over consonant duration as well as segmental duration data. The present study attempts to sort out whether the voicing adaptation mechanisms at work have phonological status by looking at the frequency of occurrence of the voicing ratios on a token-to-token basis and at the temporal location of the vocal fold vibration period, as well as at the duration of the target obstruents and of the vowel preceding the cluster. Special attention is paid to the speaker-dependent voicing adaptation strategies and to whether voicing degree is conditioned by the manner of articulation characteristics of all consonants in the cluster.

1.1 Categorical and Gradient Aspects of Regressive Voicing Assimilation

In addition to exploring voicing adaptation in a particular language (Catalan), findings reported in the present study are relevant to theoretical accounts of assimilation. A major issue in this respect is whether, based on experimental grounds, it is possible to ascertain if voicing adaptation in consonant clusters has a cognitive status and is thus assimilatory, or else should be considered a purely phonetic coarticulatory effect. It has been claimed in this respect that, if it is a phonological assimilatory process, voicing adaptation ought to be discrete and thus should not be phonetically

[1] The plural marker 's' in words such as *rams* is represented as voiceless underlyingly in spite of the lack of morphophonological alternations proving that it should be either voiceless or voiced.

Phonetica 2014;71:128–156
DOI: 10.1159/000369028

conditioned by changes in segmental context or speech rate; on the other hand, if it is a phonetic coarticulatory effect, voicing adaptation should yield a continuous range of voicing degrees and decay as a function of time (Jansen, 2004). As discussed below, however, inspection of the temporal location of voicing during the target obstruents reveals that, even if incomplete and contextually variable, voicing adaptation may be considered assimilatory and thus phonological in specific circumstances.

In principle, voicing in word-final obstruents may be maintained from consonant onset to offset either by having the vocal folds vibrate actively or else by allowing some air leakage through the nasal cavity or by expanding the oral cavity in order to compensate for an increase in intraoral pressure, which may take place in specific segmental scenarios (Westbury and Keating, 1986). In the event that word-final obstruents are partially voiced whenever speakers fail to make complete use of these voicing mechanisms, the issue remains as to whether regressive voicing applies categorically and therefore a voicing assimilation rule operates or not. Experimental work conducted on languages with stop prevoicing suggests that, if the following requirements (a) and (b) are fulfilled, regressive voicing adaptation in heterosyllabic C#C sequences may be considered to be at the same time categorical in the sense that speakers plan the word-final obstruent as voiced, and phonetically gradient since C1 voicing may be complete or partial and even absent.

(a) Voicing assimilation may be said to operate whenever there is continuous vocal fold vibration from the vowel preceding the cluster into the first cluster obstruent even if this voicing period dies out before C1 offset thus rendering the consonant partially voiced. The early presence of voicing suggests that C1 has been planned as voiced by the speaker. Data from the literature on Catalan and other languages with prevoiced stops (Catalan: Cuartero, 2001; French: Hallé and Adda-Decker, 2011; Slovak: Bárkányi and Kiss, 2014) show that, whenever present, voicing generally occurs at the left edge of C1 and may extend over the entire consonant or end before C1 offset and resume about C2 onset or somewhat later; much more rarely, voicing starts somewhere after C1 onset or is scattered throughout the consonant. Crucially, however, in order to make valid inferences about regressive voicing assimilation based on the presence of voicing at C1 onset one has to make sure that, as found for 2 Catalan speakers in a previous study on C#C voicing assimilation (Cuartero, 2001), the voicing period in question is longer than the vowel-dependent voicing lag (also known as VTT or voice termination time; see Jansen, 2004, regarding this term) which may be found in postvocalic voiceless stops. Only if voicing duration at the C1 left edge exceeds VTT duration we may thus claim that voicing adaptation in consonant clusters is assimilatory and thus related to the following underlyingly voiced consonant and not to the vowel preceding the consonant sequence (see also Hallé and Adda-Decker, 2011). In these circumstances, the failure for voicing to continue during C1 may be attributed to aerodynamic factors causing an intraoral pressure rise and the cessation of vocal fold vibration.

(b) Voicing assimilation may also be considered to apply when segmental duration is used by speakers in order to signal the target obstruent as voiced even in cases where vocal fold vibration is barely found at C1 onset. In these circumstances, word-final obstruents should be shorter and the vowel preceding the cluster longer when the consonant trigger is voiced than when it is voiceless (see Jansen, 2004, and Smith, 1997, for English). These voicing-related differences in segmental duration depend on intraoral pressure level, airflow volume and degree of closure or constriction (Kohler, 1984) and may not apply to all phonetic segments, i.e., they may operate on fricatives but not on stops (Russian: Burton and Robblee, 1997) and on the vowel rather than on the obstruent (Catalan: Strycharczuk, 2012).

In view of these considerations, the present study seeks to differentiate instances of voicing assimilation in C#C and CC#C sequences from the routine occurrence of VTT by applying the criterion defined in (a) above. In line with criterion (b), it will also be assumed that regressive voicing assimilation may be cued by the duration of word-final obstruents and/or the vowel preceding the cluster whenever the obstruents in question exhibit partial voicing.

A related goal of this investigation is to explore the extent to which Catalan speakers may differ regarding voicing degree in word-final obstruents and the voicing adaptation strategy that they use, and even whether they apply or fail to apply the voicing assimilatory process. Thus, for example, according to spectrographic data on the Greek sequences /sb, sd, sg, sl, sm/ where /s/ is supposed to assimilate to C2 in voicing, 2 speakers favored full fricative voicing, 2 speakers exhibited full or partial voicing, and 1 speaker did not seem to care about whether /s/ was realized as fully voiced, partially voiced or voiceless (Baltazani, 2006). As for Catalan, EGG data on regressive voicing assimilation in obstruent C#C sequences turned out to exhibit large differences in C1 voicing degree between the 2 speakers subject to analysis (Cuartero, 2001). This scenario resembles that for place assimilation in English /tC/ and /nC/ sequences where speakers may show regressive assimilation, partial adaptation (e.g., the tongue dorsum is raised but does not make complete velar contact during C1 = /n/ in a sequence such as /nk/) or no C1-to-C2 adaptation at all (Ellis and Hardcastle, 2002).

In a similar vein, the present study is also concerned with differences in the phonetic implementation of the regressive voicing assimilation process that may occur among languages with prevoiced stops. Thus, judging from available data on obstruent + obstruent sequences, regressive voicing appears to be close to categorical in Russian where C1 voicing ratios are often about 90% or higher (Burton and Robblee, 1997; Kulikov, 2013), and more gradient in Hungarian with ratios intermediate between those for fully voiced and voiceless consonants (Gow and Im, 2004; Markó et al., 2010). Also regarding place assimilation in /nC/ sequences, C1-to-C2 adaptation has been reported to operate almost without exception in Spanish and Italian, i.e., C1 acquires the C2 closure or constriction location throughout its entire duration in practically all sequence tokens and for all speakers (Farnetani and Busà, 1994; Celata et al., 2013), and less often in English (see above).

Rather than evaluating the obstruent voicing ratios over consonant duration, attention will be paid to the frequency distribution of different voicing ranges for all cluster repetitions. Based on a measure of frequency distribution it was concluded for French that, even though cross-token voicing ratios were less than 80%, regressive voicing adaptation in consonant clusters with a voiced C2 was categorical since voicing ratios about 90% occurred much more often (63.1% of the time) than voicing ratios below 90% (about 10% or less; Hallé and Adda-Decker, 2011). This frequency distribution measure also provides an indication of the degree of consistency in voicing implementation and presumably of articulatory control over the voicing dimension, i.e., speakers may be more or less consistent in assigning specific voicing degrees to target obstruents in C#C and CC#C sequences when data for all cluster tokens are taken into account. In particular, it could be that speakers who voice word-final obstruents most turn out to also be more consistent in their voicing behavior than those that voice them least since the former rely on glottal activity to a larger extent than the latter.

1.2 Segmental Factors Impinging on Voicing Adaptation

In addition to determining whether voicing assimilation is at work or not, the present investigation deals with the phonetic implementation of the voicing adaptation process by focusing on the role of several segmental factors: the number of word-final obstruents and their position in the C#C and CC#C sequences under analysis; whether word-final obstruents are stops or fricatives, and also whether the word-initial voicing trigger is an obstruent (a stop or a fricative) or a sonorant (a nasal, a lateral, the alveolar trill /r/ or an approximant). Other factors which may also impinge on voicing adaptation will not be looked into such as the place of articulation of the target and triggering consonants (see Recasens and Mira, 2012, 2013, in this respect) and rate of speech (Abdelli-Beruh, 2004).

As for the first conditioning factor, word-final obstruents are expected to exhibit less voicing in C#C than in CC#C sequences since an increase in the number of consonants (and in particular in the number of obstruents) should cause a rise in the intraoral pressure level and voicing to fade away during the cluster (Westbury and Keating, 1986). Moreover, voicing in CC#C sequences ought to be less during C2 than during C1 if intraoral pressure achieves its maximum towards the middle of the consonant cluster.

Regarding the effect of consonant manner of articulation on voicing degree, voicing in word-final obstruents is expected to be less whenever the target consonant is a fricative than when it is a stop in view of the difficulty involved in combining vocal fold vibration with an open glottis for the passage of considerable airflow in the case of fricatives (see Slis, 1986, for supporting evidence for Dutch). Another interesting issue is whether, in languages with obstruent voicing neutralization such as Catalan, voicing is less for word-final obstruents, which are presumably unspecified for voicing, than for syllable-initial obstruents, which are actively voiced. In line with these differences in active voicing, an increase in intraoral pressure could prevent voicing from occurring in the former obstruents rather than in the latter.

Special attention will be paid to the role of manner of articulation in the voicing assimilation trigger. Descriptive data show that languages with regressive voicing assimilation differ as to whether sonorants act as voicing triggers (Catalan, Slovak, West Flemish, Poznań and Kraków Polish, Moravian Czech) or not (Russian, Hungarian, Standard Czech) (Dvořák, 2010; Strycharczuk, 2012; Bárkányi and Kiss, 2014). In order to account for this difference, phonologists have proposed that sonorants should not act as voicing assimilation triggers because they are unspecified for voice, i.e., their production involves no voicing control since they cannot be voiceless. According to this theoretical account, voicing adaptation before sonorants results from spontaneous voicing that occurs during these consonants, mostly so in languages with final devoicing where target obstruents are supposed to become especially susceptible to voicing adaptation (Jansen and Toft, 2002; Jansen, 2004, 2007a, b). Another proposal is that sonorants are laryngeally specified as voiced in some languages but not in others (Wheeler, 2005).

An alternative explanation is that differences in obstruent voicing as a function of contextual obstruents versus sonorants are conditioned, at least in part, by production and perceptual factors. According to this view, sonorants would not necessarily behave as a class but would differ among themselves regarding the amount of voicing they induce in the preceding obstruent. Thus, for example, obstruent voicing is expected to be less before nasals, laterals and the alveolar trill than before approximants. The absence of much regressive voicing before nasals and laterals may co-occur with a delay in velar lowering and in the opening of the lateral mouth passages during the obstruent; this could be so since changes in oral pressure associated with nasality and in oral pressure and lingual configuration associated with laterality may conflict with the generation of a highly intense frication noise for fricatives and a perceptually prominent burst for stops (Solé, 2007, 2009; Ohala and Solé, 2010). As for the trill /r/, the lack of voicing in the preceding obstruent could be related to the high production requirements involved in keeping the tongue tip vibrating for a relatively long period of time, i.e., a sufficient pressure difference across the oral constriction, some tongue predorsum lowering and postdorsum retraction, and the right amount of tongue muscle tension in order to set the tongue tip into vibration (Solé, 2002). These articulatory mechanisms for the trill are to a large extent antagonistic with respect to the generation of the frication noise for fricatives and may induce an increase in intraoral pressure and thus the lack of voicing during a preceding stop. This production-based account appears to be in accordance with data for languages with regressive voicing assimilation before any voiced consonant showing that, when occurring before nasals and to some extent laterals, stops and fricatives are partially or completely voiceless and show less voicing than before voiced obstruents (Catalan: Cuartero, 2001; Recasens and Mira, 2012; Strycharczuk, 2012; Greek: Baltazani, 2006). On the other hand, to the extent that they are produced with a relatively wide constriction, approximants ought not to cause a significant increase in intraoral pressure and therefore should trigger considerable regressive voicing during a preceding obstruent. In Catalan, this scenario would apply not only to /j/ but also to the approximant realizations [β, ð, ɣ] of /b, d, g/, which may occur after fricatives in alternation with stop realizations depending on degree of constriction, speech rate and speaker. Along these lines, the failure for obstruents to assimilate in voicing to all sonorants in languages such as Russian and Hungarian could result from the phonologization of a scenario like the one found in Catalan where sonorants may trigger more or less voicing depending on their specific manner of articulation requirements.

Some support for this production-based hypothesis may also be sought in the fact that specific sonorants may not only induce less voicing than obstruents but also cause preceding word-final obstruents to exhibit considerable random voicing variability. This high voicing variability degree may be associated with a high degree of variability in the timing of those supraglottal articulatory events which are used for the coproduction of the two consecutive consonants in the cluster. Thus, for example, to the extent that the combination of frication and nasality is disfavored, the precise onset of anticipatory velar lowering in fricative + nasal sequences may turn out to vary considerably as a function of sequence token, speaker and speech rate (Solé, 2007). In order to investigate this issue, the present study will analyze possible token-to-token voicing differences for target obstruents before sonorants versus other obstruents and will try to ascertain whether these variability differences are related to differences in the acoustic duration of the word-final obstruents and, by extension, in the timing of the supraglottal articulatory events.

1.3 Summary of Research Goals

The major goal of this article is to investigate whether regressive voicing adaptation in Catalan C#C and CC#C sequences with word-final obstruents may be considered an assimilatory process and if it is phonetically categorical or gradient. Section 3.1

Phonetica 2014;71:128–156
DOI: 10.1159/000369028

deals with differences in obstruent voicing in C#C sequences as a function of speaker, token and the manner of articulation characteristics of both the target consonant and the word-initial voicing trigger. In an effort to determine whether voicing adaptation is assimilatory or not, section 3.2 investigates if voicing at the left edge of C1 in C#C sequences is longer than vowel-related VTT and the extent to which segmental duration cues voicing assimilation. Section 3.3 deals with identical research issues for CC#C sequences.

2 Method

Regressive voicing adaptation was analyzed using vocal fold vibration data for the following 60 C#C and 148 CC#C combinations: (C#C) /p, t, k, f, s, ʃ/ + /b, d, g, z, m, n, l, r, ʎ, j/; (CC#C with a fricative C2) /ps, ts, ks, fs, ls, rs, ms, tʃ, mf, lf, rf/ + /b, d, g, m, n, l, ʎ, j/; (CC#C with a stop C2) /lp, rp, sp, lk, rk, sk/ + /b, d, g, m, n, l, z, r, ʎ, j/. These consonant clusters were embedded in meaningful sentences which were about 7/8 syllables long (e.g., /fr/ in the sentence *a la cuina hi ha un xe**f r**os* 'there is a blond chef at the kitchen') and bore a sentence stress on the vowel following the target cluster.

EGG and acoustic data were recorded simultaneously by 8 middle-aged native Catalan speakers, i.e., 5 women (E.V., M.A., P.E., L.O., V.A.) and 3 men (S.O., M.O., D.R.), with the EGG-2 glottograph from Glottal Enterprises and the multichannel Kay Pentax system. These informants came from different areas of Catalonia: 6 of them speak the Eastern Catalan dialect and were born in urban Barcelona (S.O., P.E.) and in other towns and villages (M.O., Banyoles; L.O., Montblanc; D.R., Tarragona; V.A., Cadaqués); the remaining 2 subjects speak Western Catalan and were born in the Baix Urgell region (E.V., M.A.). All cluster productions were expected to conform to the regressive voicing assimilation rule independently of the speakers' dialectal origin (no differences in cluster voicing assimilation among Catalan dialects have been noted in the literature). Sentences were read 8–10 times at the speakers' normal speech rate, and 7 rather than all 10 cluster tokens were chosen for analysis so as to avoid having to process an unmanageable number of items. Overall, 11,648 cluster tokens were analyzed (3,360 C#C tokens + 8,288 CC#C tokens). All signals were acquired at 44,100 Hz, and the EGG signal was downsampled to 500 Hz, and the acoustic signal to 11,025 Hz. The EGG signal was smoothed and analyzed using the MatLab script Peakdet 2 (Abadal and Recasens, 2009).

V, C1 and C2 onsets and offsets were estimated based on visual inspection of simultaneous spectrographic and waveform displays using the CSL (Computer Speech Lab) analysis program from Kay Pentax. Phonetic segments were delimited by the edges of a period of high-intensity formant structure for the vowels preceding and following the consonant cluster, 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, productions of word-initial /b, d, g/ after a fricative were classified as [β, ð, ɣ] (approximants) if exhibiting weak formants occasionally with some frication noise overimposed, or as [b, d, g] (voiced stops) if showing no formant structure and generally a burst. The alveolar rhotic was usually identified by the presence of one or more short closures (it shows regularly more than one contact word-initially where it is typically realized as a trill); if realized sometimes as an approximant or a fricative, the onset and offset of /r/ were determined applying the same criteria for /b, d, g/ above.

Peakdet 2 was used in order to find out whether a particular consonant exhibited continuous or partial voicing. In the latter event, the onset or offset of voicing after or before a period of voicelessness, respectively, were identified at the time at which the positive peak of the first derivative of the relevant glottal pulses occurred. A threshold detection at 25% of the positive maximum of the first derivative was applied, which is slightly below other threshold values proposed in the literature (Rothenberg and Mahshie, 1988). The voicing measures for the word-final obstruents were the voicing ratio over consonant duration and the frequency of occurrence of several voicing ranges (0–20, 20–40, 40–60, 60–80, 80–100%). Moreover, in order to explore differences in the degree of C1 voicing variability among consonants in C#C sequences, coefficients of variation for the C1 voicing ratios were obtained for each consonant and each speaker in each C2 condition by dividing

DOI: 10.1159/000369028

the standard deviation values across tokens by the corresponding mean. No data were processed for those C#C sequences where C1 shares the same place and manner of articulation as C2 (/pb, td, kg, sz/, and also /ʃz/, which is realized [zʲ(:)] in Catalan), and for those CC#C sequences where a stop C2 was absent on the spectrographic displays (40% of tokens of /spC, skC/, 10% of tokens of /lpC, lkC, rpC, rkC/).

Positive VTT values were measured at C1 onset of the voiceless clusters /pk, tk, kt, fk, sk, ʃk/ embedded in 6 short sentences (e.g., /tk/ in the sentence *passa un solda**t** **c**urd* ‘a Kurdish soldier goes by’). Data were collected for 7 cluster tokens produced by all 8 Catalan subjects and thus 336 cluster tokens overall.

In order to investigate whether segmental duration contributed to signaling regressive voicing adaptation, the duration of the word-final obstruents and the vowel preceding the cluster was analyzed for 10 C#C and 4 CC#C sequence pairs differing in the voiced or voiceless status of the word-initial consonant: (stop + stop sequences) /pt/ (as in *no queda ca**p** **t**alp* ‘there is no mole left’) -/pd/ (*no queda ca**p** **d**au* ‘there is no die left’), /pk/-/pg/, /kp/-/kb/, /kt/-/kd/; (stop + fricative) /ps/-/pz/, /ts/-/tz/, /ks/-/kz/; (fricative + stop) /sp/-/sb/, /st/-/sd/, /sk/-/sg/; (stop + fricative + stop) /psp/-/psb/, /tst/-/tsd/, /ksk/-/ksg/; (sonorant + stop + fricative) /lps/-/lpz/. These clusters were produced 10 times by all 8 speakers (the C#C pairs) and by subjects D.R., E.V., M.O. and S.O. (the CC#C pairs) in meaningful sentences showing the same stress pattern and number of syllables as the utterances used for carrying out the voicing measures. Segmental duration for C1 and C2, or for C1, C2 and C3, and for the vowel preceding the cluster were measured for all cluster tokens applying the same segmentation criteria as above. Overall data for 1,920 cluster tokens were processed (1,600 for the C#C sequences, 320 for the CC#C clusters).

Several General Linear Model analyses were run with SAS version 9.3 on the voicing data with ‘speaker’ as a fixed factor so as to detect significant differences among individual speakers and significant interactions between this and other variables. The following tests were carried out in order to uncover segmental effects on word-final obstruent voicing: two tests on the voicing ratios and cross-token coefficients of variation for C1 in C#C sequences, with ‘C1’, ‘C2’ and ‘speaker’ as fixed variables; two tests on the C1 voicing ratios for the CC#C sequences /psC, tsC, ksC/ (with the fixed factors ‘C1’, ‘C3’, ‘speaker’) and /spC, skC/ (with the fixed factors ‘C2’, ‘C3’, ‘speaker’); two other tests on the C2 voicing ratios for the CC#C sequences /psC, tsC, ksC, msC, lsC, rsC, fsC/ (with the fixed factors ‘C1’, ‘C3’, ‘speaker’) and /lpC, lkC, rpC, skC, spC, skC/ (with the fixed factors ‘C1’, ‘C2’, ‘C3’, ‘speaker’). Two additional General Linear Model analyses were performed on VTT and the duration of the voicing period occurring at C1 onset in C#C and CC#C sequences, with the fixed factors ‘C1’, ‘condition’ (with the variable levels ‘stop C1’ and ‘fricative C1’) and ‘speaker’. Finally, separate tests were run on the duration of C1 and the vowel preceding the cluster in C#C sequences and of C1, C2 and the vowel preceding the cluster in CC#C sequences, with ‘syllable-initial consonant’, ‘C1’ or ‘C1’ and ‘C2’, and ‘speaker’ as fixed variables. Whenever applicable, pairwise comparisons among variable levels were carried out using the Bonferroni correction. Results for the main effects and two- and three-factor interactions at the $p < 0.001$ level of significance will be reported.

3 Results

3.1 Voicing Variability in C#C Sequences

3.1.1 Speaker-Dependent Voicing Differences

Catalan speakers are clearly divided regarding the extent to which C1 adapts to C2 in voicing in C#C sequences, as determined by the period of vocal fold vibration. Indeed, C1 voicing ratios over C1 duration yielded a main effect of speaker [$F(7, 2{,}933) = 178.84$, $p < 0.001$] which according to post hoc tests was related to a decrease in voicing degree in the progression D.R., V.A. (mean values of 76.1 and 72.7%) > L.O. (60.7%) > M.O. (51.7%) > P.E. (45%) > E.V., M.A. (40, 37.7%) > S.O. (19.3%).

As pointed out in section 1.1, an informative way to represent the speaker-dependent differences in C1 voicing adaptation is by plotting the frequency distribution of several C1 voicing ranges across C2 conditions, i.e., maximal voicing when the C1 voicing ratio amounts to 80% or higher (80–100%), minimal voicing (0–20%) and intermediate or partial voicing (20–40, 40–60, 60–80%). As shown in figure 1, the frequency distribution of these five C1 voicing ranges allows classifying speakers into two major groups depending on whether they favor maximal C1 voicing (speakers V.A., D.R. and L.O., fig. 1a) or minimal C1 voicing (speakers P.E., M.A., E.V. and S.O., fig. 1b). These two groups of speakers differ considerably regarding the frequency of occurrence of the maximal and minimal voicing ranges: 80–100% voicing takes place about 40–60% of the time for speakers who voice most and less than 25% for those who voice least, and the opposite situation applies to the 0–20% voicing range. There also are small voicing differences among speakers within each group: speakers V.A. and D.R. show more C1 voicing than L.O. (fig. 1a), and speaker S.O. exhibits less voicing than P.E., M.A. and E.V. (fig. 1b). Moreover, the fact that the maximal 80–100% voicing range does not occur 80% of the time or higher in the case of subjects V.A., D.R. and L.O. indicates that C1 voicing adaptation for speakers who voice most applies gradiently and is thus subject to severe contextual restrictions. The figure also reveals a trend for the intermediate 20–40, 40–60 and 60–80% voicing ranges to occur not more often than 25% of the time for the two groups of speakers, which suggests that we may be facing a categorical scenario with subjects who assimilate and subjects who do not assimilate. Speaker M.O. (fig. 1c) exhibits similar voicing percentages for most voicing ranges and thus lies between the two major subject groups regarding the extent to which C1 adapts in voicing to C2. In view of the special behavior of this speaker, his data will only be referred to whenever relevant to the issue being analyzed.

In order to gain some understanding about why speakers may differ so remarkably in C1 voicing degree, a correlation analysis was performed on their mean C1 and C2 voicing ratios across segmental conditions. Correlation values for all clusters, and separately for those with a stop C1 and for those with a fricative C1, turned out to be quite high, i.e., $r = 0.75–0.80$, thus meaning that, as shown by figure 2, subjects show more or less voicing during C1 depending on whether they exhibit more or less voicing during C2. This finding points to differences in the extent to which the Catalan speakers subject to investigation voice consonants in general and not only word-final obstruents before a voiced consonant in clusters.

3.1.2 Consonant Production Requirements

C1 voicing in C#C sequences turned out to be strongly conditioned by the articulatory characteristics of both C1 (the target consonant) and C2 (the voicing trigger), which supports the notion that regressive voicing adaptation is implemented gradiently in heterosyllabic consonant clusters in Catalan.

Regarding the effect of C1, fricatives were found to exhibit about 20% less voicing than stops with overall means of 40.7 and 59.9%, respectively [$F(1, 2,933) = 360.47$, $p < 0.001$]. Moreover, there was a speaker × C1 interaction [$F(7, 2,933) = 44.24$, $p < 0.001$] according to which this difference was significant for the speakers V.A., D.R., L.O., P.E. and M.A. but not for M.O. ad S.O., while speaker E.V. showed more voicing for fricatives than for stops. The finding that fricatives allow less regressive voicing than stops indicates that vocal fold vibration during C1 is harder to maintain when

DOI: 10.1159/000369028

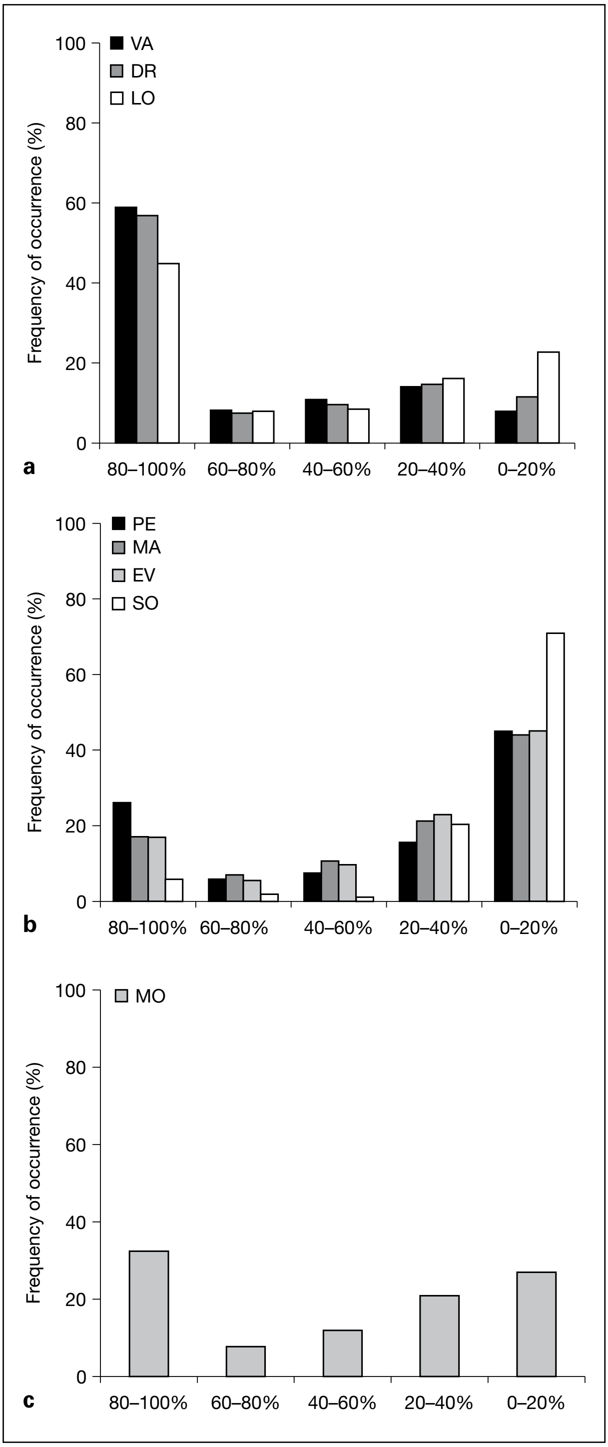

Fig. 1. Frequency of occurrence of C1 voicing ranges in C#C sequences. Data are plotted for three groups of speakers showing different C1 voicing degrees (**a–c**).

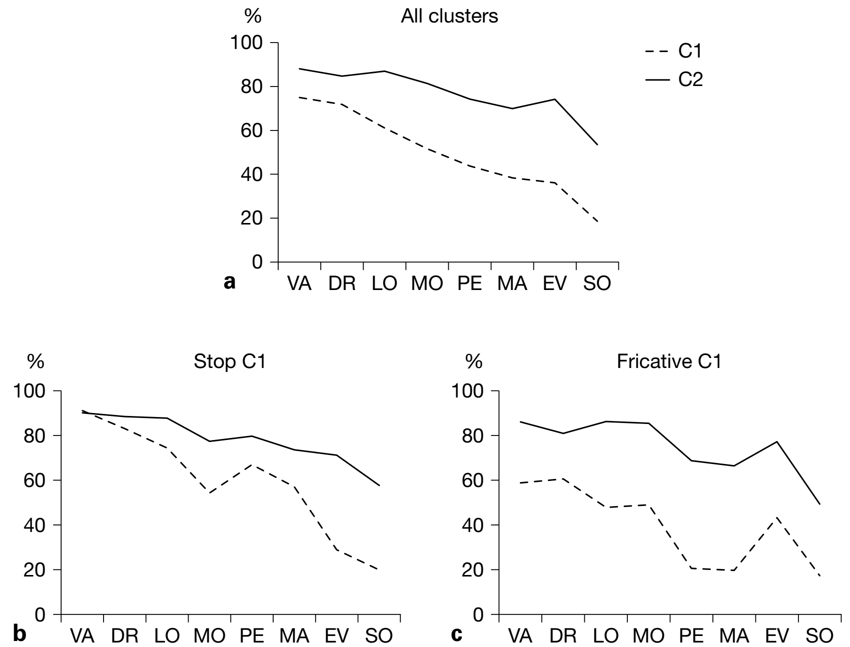

Fig. 2. Speaker-dependent voicing ratios for C1 and C2 in C#C sequences. Data are presented across consonant conditions (**a**) and for clusters with a stop C1 and for those with a fricative C1 (**b**, **c**). C1 voicing values are given from highest to lowest as appearing in **a**.

the glottis stays relatively open for the passage of considerable airflow. Moreover, C1 voicing turned out to be negatively correlated with duration across speakers ($r = –0.63$), with fricatives being longer and less voiced than stops. Data on the frequency distribution of the five voicing ranges for word-final stops and fricatives plotted in figure 3 reveal that maximal voicing occurs 60–80% of the time for stops in the case of speakers who voice most, which approaches complete voicing adaptation (fig. 3a), and less than 20% of the time for fricatives in the case of speakers who voice least thus indicating that these consonant productions are practically voiceless (fig. 3d). All voicing ranges exhibit a more even frequency of occurrence in the two remaining scenarios, i.e., in the case of sequences with a fricative C1 for V.A., D.R. and L.O. and of those with a stop C1 for the remaining subjects (fig. 3b, c).

There was also a considerable effect of the C2 manner of articulation on C1 voicing. The statistical analysis yielded a C2-dependent significant effect [$F(9, 2{,}933) = 44.82$, $p < 0.001$], which according to post hoc tests was associated with more C1 voicing before obstruents and /j/ than before nasals, laterals and the alveolar trill (also before /b/ than before /d, g, z, j/). These C1 voicing differences may be traced in the two graphs of figure 4 showing the percentages of occurrence of the C1 voicing ranges as a function of the five C2 manners of articulation for speakers who show most voicing (fig. 4a) and for those who show least voicing (fig. 4b). According to the former group of subjects, obstruent + obstruent sequences favor maximal C1 voicing about

DOI: 10.1159/000369028

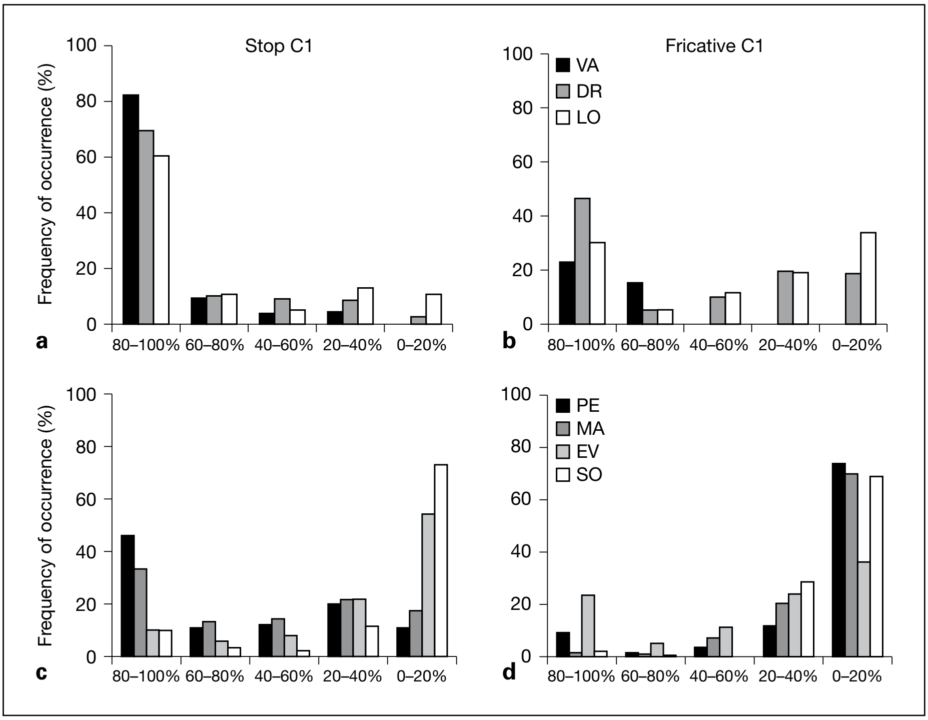

Fig. 3. Frequency of occurrence of voicing ranges for word-final stops (**a**, **c**) and fricatives (**b**, **d**) in C#C sequences. Data are plotted for two groups of speakers showing different voicing degrees.

70% of the time, which suggests that regressive voicing assimilation could be at work in this case. On the other hand, for the two groups of speakers, sonorants do not behave uniformly regarding C1 voicing adaptation: there is much less C1 voicing before a nasal, a lateral or the alveolar trill than before /j/, and as much C1 voicing before /j/ as before an obstruent. Statistical results also yielded a speaker × C2 interaction [$F(63, 2{,}933) = 4.12$, $p < 0.001$] according to which, while most speakers favor voicing before obstruents and /j/ (M.A., V.A., E.V., L.O.), other subjects do so only before /z/ (S.O.), before /z, j/ (P.E.) or before obstruents but not before /j/ (D.R.).

Two significant C1 × C2 and speaker × C1 × C2 interactions [$F(9, 2{,}933) = 6.71$, $p < 0.001$; $F(63, 2{,}933) = 5.37$, $p < 0.001$] reveal that the extent to which the C2 manner of articulation triggers more or less voicing during C1 depends on whether C1 is a stop or a fricative. Indeed, according to the frequency of occurrence data plotted in figure 5a, b, nasals and to a lesser extent laterals and the alveolar trill induce comparatively less voicing in a preceding fricative than in a preceding stop; thus, for example, 80–100% voicing takes place about 70% of the time in stop + nasal sequences and only about 10% in fricative + nasal sequences. Figure 5a, b also shows that word-initial obstruents and /j/ trigger more voicing than nasals, laterals and /r/ independently of whether C1 is a stop or a fricative. As for the speakers exhibiting least voicing, figure 5c, d exhibits similar C2-dependent voicing differences during word-final stops and fricatives to those just described. C1 voicing degree was also

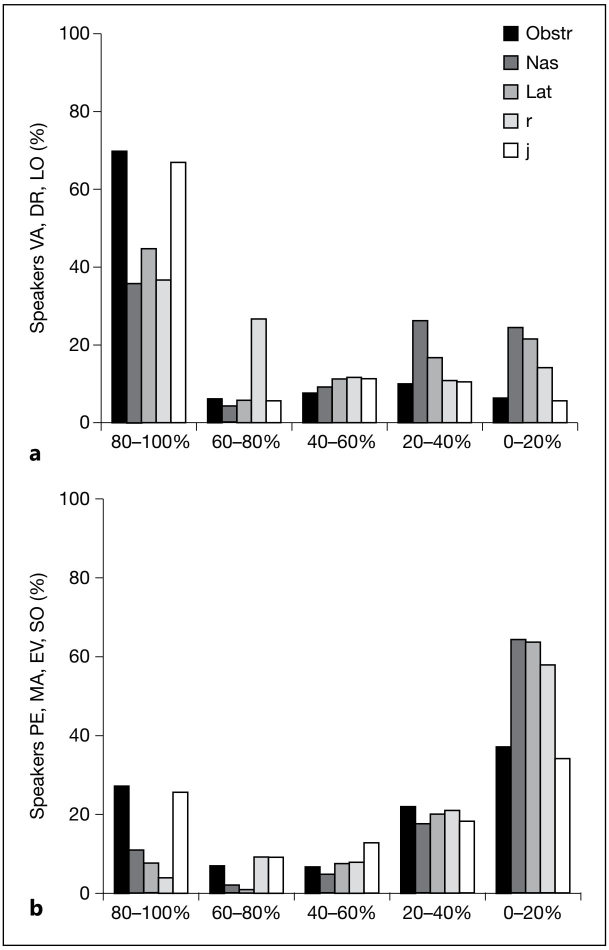

Fig. 4. Frequency of occurrence of C1 voicing ranges as a function of C2 in C#C sequences. C2 has been split into five manner classes, i.e., obstruents, nasals, laterals, the trill /r/ and /j/. Data are plotted for two groups of speakers showing different voicing degrees (**a**, **b**).

conditioned by whether syllable-initial /b, d, g/ were realized as stops or as approximants when occurring after a fricative (see section 1.2). The effect of this manner of articulation difference becomes apparent in figure 6, which plots voicing ratios for a fricative C1 against those for stop or approximant realizations of C2 = /b, d, g/; indeed, according to the circles and fitting lines displayed in the figure, fricative voicing ratios are higher before [β, ð, ɣ] than before [b, d, g] [$F(1, 219) = 24.57$, $p < 0.001$]. In sum, data reported in figures 4–6 reveal that approximants should not be grouped with nasals, laterals and rhotics regarding the degree of voicing that they induce in preceding word-final obstruents.

To recapitulate, C1 voicing turned out to be less for fricatives than for stops in view of the conflicting aerodynamic and voicing demands involved in the production of the former consonants. Other conflicting production requirements referred to in section 1.2 may also explain why C1 voicing (mostly if the consonant is a fricative) is less before nasals, laterals and the alveolar trill than before /j/ and stops (mostly if realized as approximants).

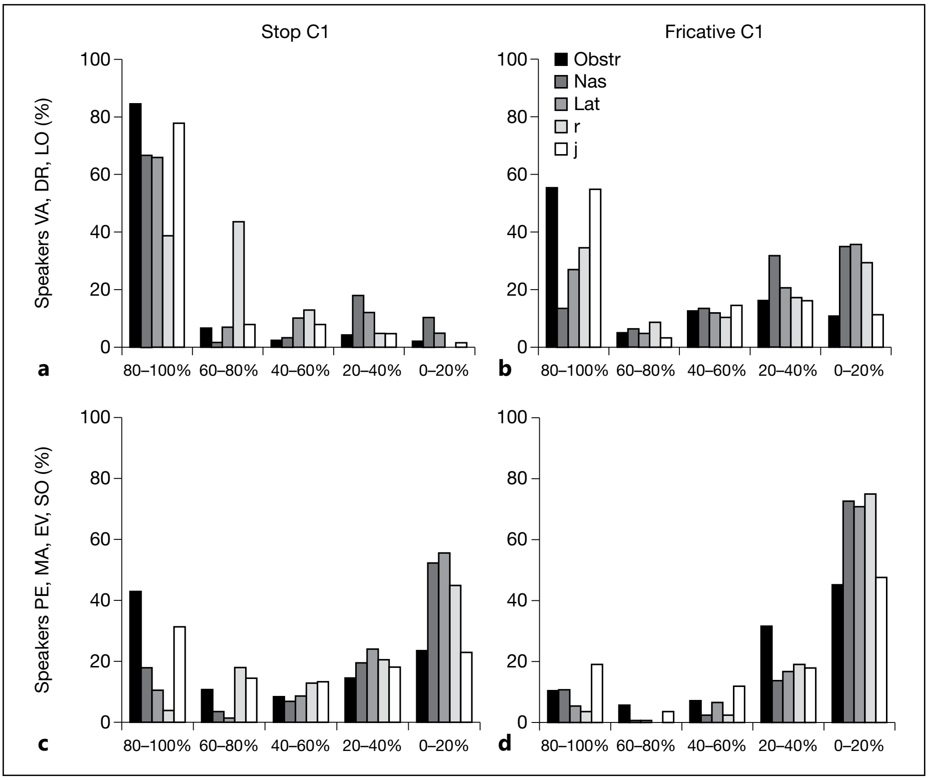

Fig. 5. Frequency of occurrence of voicing ranges for word-final stops (**a**, **c**) and fricatives (**b**, **d**) as a function of C2 in C#C sequences. C2 has been split into five manner classes, i.e., obstruents, nasals, laterals, the trill /r/ and /j/. Data are plotted for two groups of speakers showing different voicing degrees.

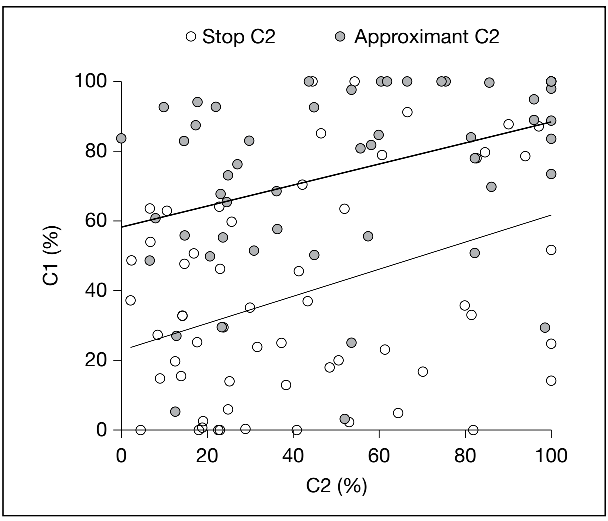

Fig. 6. Cross-token voicing ratios for C1 and C2 in fricative + /b, d, g/ sequences plotted as a function of whether C2 is realized as a stop or as an approximant. A best-fitting straight line has been appended to the two C2 data sets.

3.1.3 Token-to-Token Variability

As stated in section 1.1, token-dependent C1 voicing variability could be indicative of the degree of articulatory control that speakers exert over the vocal folds for implementing regressive voicing adaptation in consonant clusters: the more consistent the C1 voicing ratios across cluster tokens, the higher the degree of voicing control.

Results for the statistical analysis run on the coefficient of variation values yielded a main speaker effect [$F(7, 63) = 19.38$, $p < 0.001$] which turned out to be related to more token-to-token variability for speakers who voice C1 least than for those who voice C1 most. Coefficient of variation values were highest for subjects S.O. (98.6) and E.V. (82.8), lowest for speakers D.R. (43.7) and V.A. (30.7), and intermediate for speakers showing intermediate C1 voicing, i.e., P.E. (73.0), M.O. (68.5), L.O. (59.5) and M.A. (52.2) (see section 3.1.1 regarding speaker-dependent differences in C1 voicing).

C1 voicing variability across tokens was also conditioned by the phonetic characteristics of C1 and C2. A main C1 effect was associated with more variability for fricatives than for stops [70.6 vs. 56.6; $F(1, 63) = 16.08$, $p < 0.001$] which, according to a significant speaker × C1 interaction [$F(7, 63) = 11.80$, $p < 0.001$], turned out to hold for those subjects who showed most C1 voicing, i.e., V.A., D.R. and L.O., and for the speakers P.E. and M.O. as well. Coefficient of variation values were also affected by C2 [$F(9, 63) = 4.23$, $p <0.001$], i.e., they were higher when C1 was followed by most sonorants (nasals, laterals, /r/) and /b/ than by most obstruents (/d, g, z/) and /j/, though these C2-dependent differences achieved significance in a small number of cases. This inverse relationship between voicing degree and voicing variability may be observed by comparing the mean C1 voicing ratios as a function of C2 (fig. 7a) with the corresponding cross-token coefficients of variation (fig. 7b). Overall, C1 voicing appears to be subject to lesser articulatory control by speakers whenever showing least voicing because of its inherent phonetic characteristics (as in the case of target fricatives) or the influence from specific contextual segments (as for obstruents before a subset of sonorants).

As hypothesized in section 1.2, random variability in C1 voicing implementation could be higher in those C#C scenarios where it may be especially hard to combine specific supraglottal events for the two consecutive consonants in the cluster such as those responsible for C1 frication and for C2 nasality, laterality or trilling. If so, those consonant sequences showing more random C1 voicing variability should also exhibit more variability in C1 duration. In order to explore this issue, figure 7c, d displays C1 duration values as a function of C2 (fig. 7c) and the corresponding cross-token coefficients of variation (fig. 7d). Figure 7b, d reveals a fairly good relationship between the correlation coefficients ($r = 0.62$) for the C1 voicing ratios and C1 duration values across C2 conditions and speakers, with C1 being most variable before nasals where, as shown by figure 7a, c, the consonant shows less voicing and is shorter than in the other C2 conditions. Therefore, more random voicing variability for obstruent + nasal sequences than for the remaining clusters appears to co-occur with more random variability in C1 duration and presumably in the onset of velar lowering. A similar relationship could apply to obstruent + lateral and obstruent + /r/ sequences: indeed, C1 voicing and duration variability across cluster tokens are often higher for these sequences than for obstruent + obstruent and obstruent + /j/ combinations.

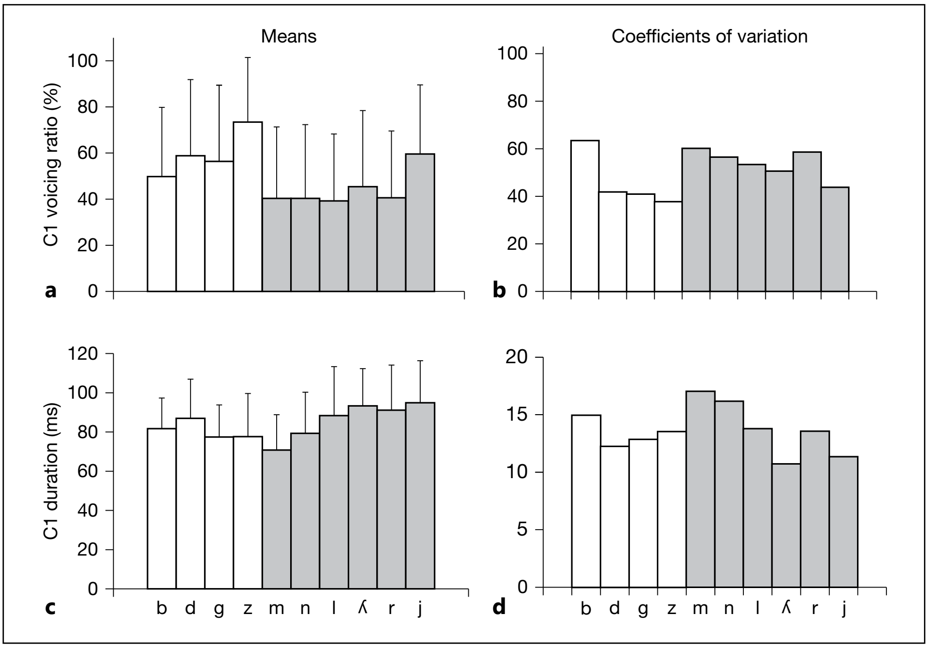

Fig. 7. Mean C1 voicing ratios and C1 duration values for C#C sequences plotted as a function of different C2 (**a**, **c**), and corresponding coefficient of variation values across tokens (**b**, **d**). Unfilled bars correspond to obstruents and filled bars to sonorants. Error bars represent one standard deviation from the mean.

3.2 Voicing Assimilation Cues in C#C Sequences

3.2.1 C1 Voicing and VTT

The preceding sections show that C1 voicing in C#C sequences as cued by vocal fold vibration varies quite considerably as a function of speaker, token and the production characteristics of C1 and C2. The issue remains as to whether, in spite of the gradient nature of the voicing adaptation process and as predicted by the Catalan assimilation voicing rule, the speakers subjected to analysis in the present investigation assimilate C1 to C2 in voicing. The following requirement (referred to already in section 1.1) should be fulfilled in order for C1 voicing adaptation to be assigned an assimilatory rather than a purely coarticulatory status: a voicing period longer than VTT ought to occur at the left edge of C1. The failure for voicing to be present at C1 onset or for voicing at the C1 left edge to be shorter than VTT implies that C1 is not planned as voiced.

In order to investigate this issue, we computed for those cases where C1 exhibited partial voicing the number of tokens where voicing occurred continuously from the vowel into C1, and compared the duration of this C1 voicing period to VTT duration. This operation was carried out separately for each speaker and each of the 6 obstruents which may appear in C1 position (see section 2 regarding the voiceless C#C sequences where the VTT data was gathered from). Cross-speaker data presented in table 1 reveal that the voicing period at C1 onset accounts for 83.2–96.6% of the overall C1 voicing duration (top) and is consistently longer than VTT (middle). The bottom panels

Table 1. Measures relating the voicing period at C1 onset, VTT and C1 voicing and duration in C#C sequences

	Voicing at C1 onset/C1 voicing ratios, %			
	mean	SD		
p	91.5	6.16		
t	96.6	3.35		
k	95.2	5.00		
f	90.4	7.54		
s	83.2	13.67		
ʃ	84.3	6.00		
	Voicing at C1 onset, ms		**VTT, ms**	
	mean	SD	mean	SD
p	37.3	22.35	8.1	5.56
t	35.3	20.60	7.1	4.62
k	23.2	9.89	15.0	11.03
f	16.6	5.09	6.4	2.82
s	18.5	7.50	6.3	3.99
ʃ	24.2	6.69	9.9	5.16
	Voicing at C1 onset/C1 duration ratios, %		**VTT/C1 duration ratios**	
	mean	SD	mean	SD
p	42.8	21.81	11.0	8.51
t	45.0	21.64	10.2	7.25
k	32.9	13.67	21.0	16.19
f	19.5	7.46	7.2	3.29
s	21.5	8.92	6.5	3.99
ʃ	25.4	7.83	9.0	4.39

Data presented in the left half correspond to clusters with a voiced C2 and those presented in the right half to clusters with a voiceless C2.

in the table show that voicing at the C1 left edge is also longer than VTT when the corresponding voicing ratios over C1 duration are taken into account [$F(1, 2{,}378) = 396.45$, $p < 0.001$]. Overall, the VTT ratios presented in the bottom right panel of the table are similar to those reported for Catalan stop, fricative + voiceless stop sequences in another voicing assimilation study (Cuartero, 2001) and for /sC/ clusters in Mexican Spanish (Schmidt and Willis, 2011), and shorter than those reported for French (Hallé and Adda-Decker, 2011). Moreover, according to two significant C1 × condition and speaker × condition interactions [$F(5, 2{,}378) = 15.87$, $p < 0.001$; $F(7, 2{,}378) = 8.24$, $p < 0.001$] and as shown by figure 8, voicing at C1 onset turned out to be longer than VTT when C1 was a stop than when it was a fricative and for all speakers except for the subject who exhibited least C1 voicing (S.O.).

The following step was to compute the number of cluster tokens where C1 was considered to be planned as voiced whether because it showed 100% voicing or else partial voicing and a voicing period at its left edge which was longer than VTT. Tokens with 0% voicing and those with an initial voicing period whose duration was equal to or less than VTT were excluded from the computation procedure. Figure 9 plots the

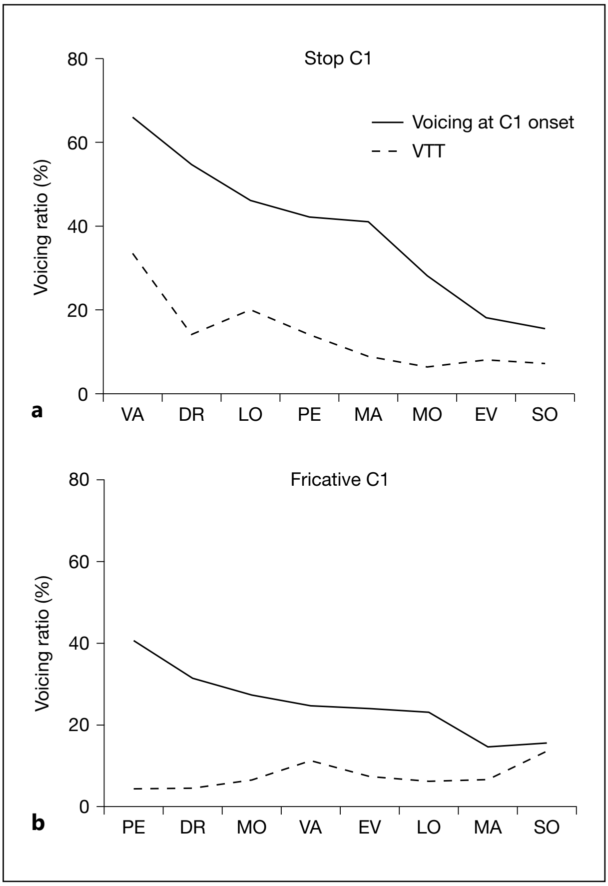

Fig. 8. Speaker-dependent voicing ratios for the voicing period at C1 onset in C#C sequences compared with the VTT values. Data are plotted separately for clusters with a stop C1 (**a**) and with a fricative C1 (**b**). In the two graphs, speaker-dependent voicing values at C1 onset are given from highest to lowest.

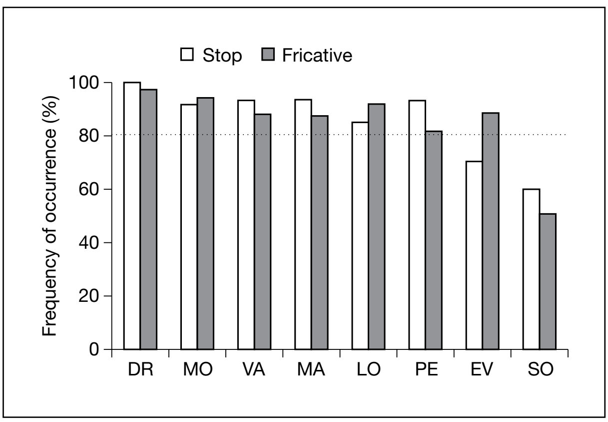

Fig. 9. Frequency of occurrence of estimated regressive assimilation cases in C#C sequences plotted as a function of C1 manner of articulation and speaker. A no-assimilation baseline has been inserted at the 80% frequency of occurrence.

Phonetica 2014;71:128–156
DOI: 10.1159/000369028

frequency of occurrence of the assimilated C1 productions over the total number of cluster tokens subject to analysis, as determined by the above criterion, as a function of speaker and of whether C1 is a stop or a fricative. If we assume that only speakers with 80% or more cluster tokens targeted as voiced have a voicing assimilation rule, the height of the bars indicates that regressive voicing assimilation is at work for speakers D.R., M.O., V.A., M.A., L.O. and P.E., but not for speaker S.O. Regarding speaker E.V., voicing for fricatives though not stops appears to comply with the assimilation criterion.

To recapitulate, most Catalan speakers subject to analysis in the present investigation appear to assimilate C1 to C2 in voicing while showing different degrees of C1 voicing adaptation depending on the articulatory and/or aerodynamic requirements for the consonants in the cluster.

3.2.2 Segmental Duration

Additional information about whether regressive voicing assimilation is at work in Catalan may also be sought by looking into other phonetic characteristics besides glottal activity such as C1 and preceding vowel duration (see section 1.2). As for consonant duration, C1 turned out to be significantly longer if followed by a voiceless consonant than by a voiced one independently of whether it was a stop or a fricative [93.8 vs. 83.9 ms; $F(1, 1{,}096) = 37.99$, $p < 0.001$]; moreover, according to a significant C1 × C2 interaction [$F(1, 1{,}096) = 15.09$, $p < 0.001$], this difference occurred in stop + stop and fricative + stop sequences, but not in stop + fricative clusters where C1 was shorter, not longer before a voiceless than a voiced C2. On the other hand, as expected, the vowel preceding the cluster was longer in clusters with a voiced C2 than in those with a voiceless C2 [75.0 vs. 66.2 ms; $F(1, 1{,}120) = 226.32$, $p < 0.001$]. Moreover, the fact that there was no significant speaker × C2 interaction for any of the C1 and vowel duration measures of interest indicates that vowel and C1 duration plays an active role in marking the presence versus absence of regressive voicing assimilation for all Catalan speakers. This includes speakers S.O. and E.V., who are reluctant to assimilate C1 to C2 in voicing when the vocal fold vibration data are taken into account. Figure 10 shows that these 2 subjects use C1 and vowel duration as C1 voicing cues in the case of most C#C pairs subject to analysis. Indeed, C1 is longer before a voiceless C2 than before a voiced C2 in stop + stop and fricative + stop sequences (compare the filled and unfilled bars in fig. 10b), and the vowel preceding the cluster is often longer if C2 is voiced than if it is voiceless, mostly so in fricative + stop sequences (fig. 10a).

The vocal fold vibration and segmental duration data reported so far allow concluding that all 8 Catalan speakers apply a regressive voicing assimilation process to the C#C clusters under analysis. Vocal fold vibration and segmental duration are used for identifying word-final obstruents as voiced when occurring before a voiced C2 by those speakers who voice most (V.A., D.R., L.O.) and by those who exhibit intermediate C1 voicing (M.O., M.A., P.E.), while segmental duration appears to play a major role in the case of those subjects who show minimal C1 voicing (E.V., S.O.).

3.3 CC#C Sequences

Another goal of the present investigation was to find out whether, as claimed by descriptive accounts, regressive voicing assimilation operates not only on an immediately adjacent obstruent (C1 in C#C sequences, C2 in CC#C sequences) but also on a

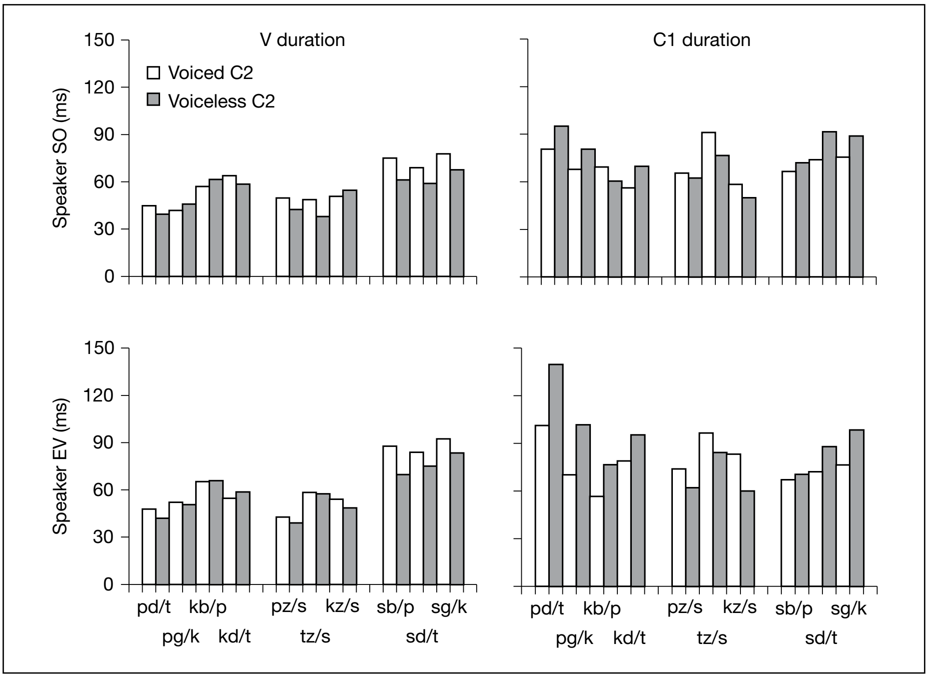

Fig. 10. C1 and preceding vowel duration for stop + stop, stop + fricative and fricative + stop cluster pairs according to speakers S.O. and E.V.

distant obstruent (C1 in CC#C sequences). Research on CC#C sequences is crucial in order to ascertain whether speakers plan word-final obstruents as voiced when occurring two segments in advance of the voicing trigger.

Except for C1 voicing for /s/ before /p, k/ in the sequences /spC, skC/, voicing ratios for obstruents in C1 and C2 position in CC#C sequences yielded a main speaker effect [C1 for /psC, tsC, ksC/, $F(7, 737) = 13.03$, $p < 0.001$; C2 for /psC, tsC, ksC, msC, lsC, rsC, fsC/, $F(7, 1{,}984) = 121$, $p < 0.001$; C2 for /lpC, lkC, rpC, rkC, spC, skC/, $F(7, 1{,}984) = 120.54$, $p < 0.001$]. Post hoc tests revealed that speaker-dependent differences in CC#C sequences are analogous to those occurring in C#C sequences: data on the frequency of occurrence of the C1 and C2 voicing ranges plotted in figure 11 show indeed maximal voicing for V.A., D.R. and L.O. (top), minimal voicing for P.E., M.A., E.V. and S.O. (middle) and intermediate voicing for speaker M.O. (bottom). There are other speaker-dependent trends. According to the data for V.A., D.R. and L.O., consonant voicing is more or less evenly distributed across most voicing ranges except for C2 = /p, k/ in the case of speaker V.A., where 80–100% voicing occurs about 70% of the time, and for C1 = /s/ in the case of speakers D.R. and L.O., which shows practically no voicing. On the other hand, C1 and C2 are mostly voiceless for subjects P.E., M.A., E.V. and S.O. and to a large extent for speaker M.O. as well. Segmental effects are analogous to those operating in C#C sequences: voicing was less for target fricatives than for target stops (compare the graphs for C1 = /p, t, k/ and C1 = /s/ and those for C2 = /p, k/ and C2 = /s/

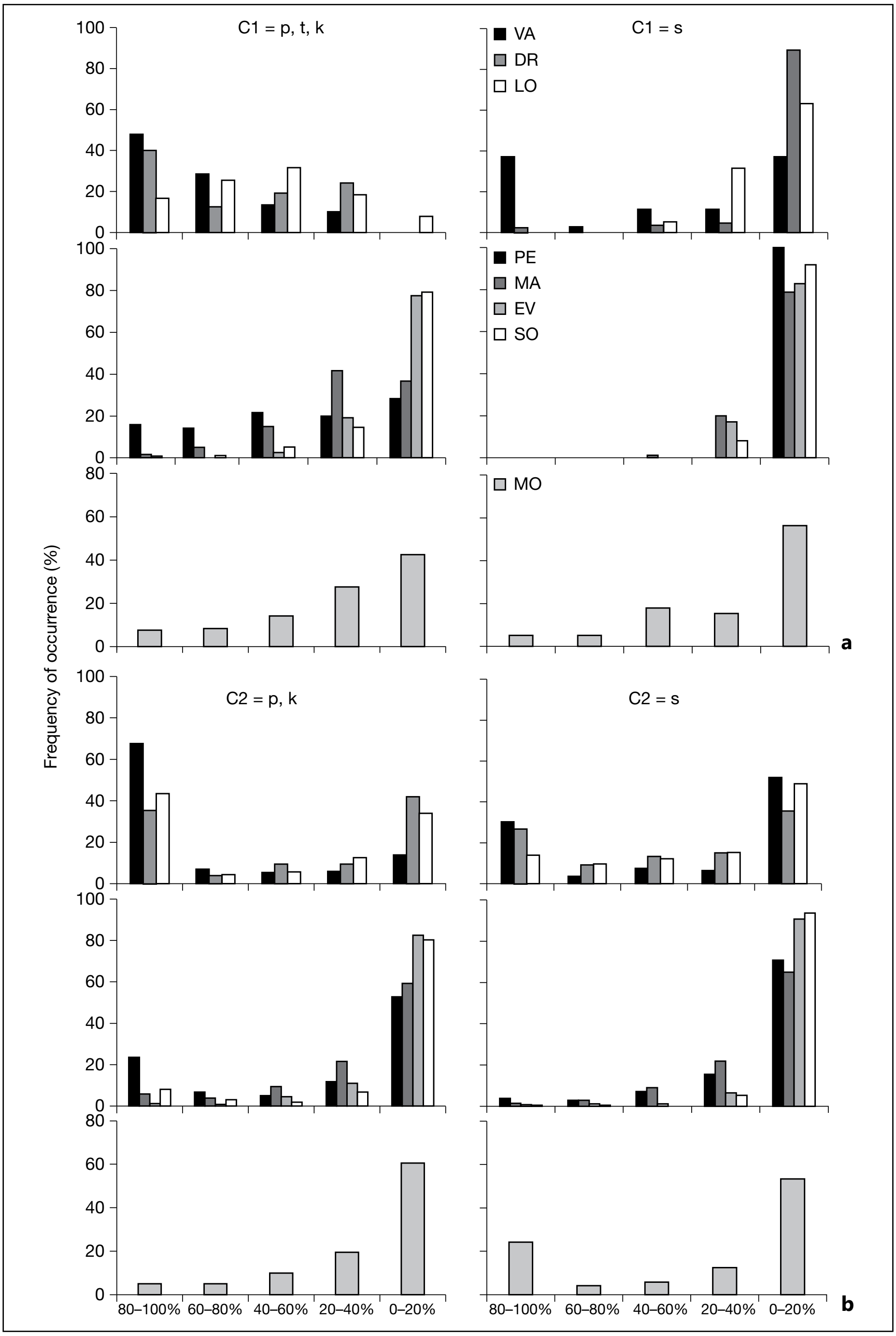

Fig. 11. Frequency of occurrence of voicing ranges for stops and fricatives occupying the C1 (**a**) and C2 (**b**) positions in CC#C sequences. Data have been plotted for three groups of speakers showing different voicing degrees.

Table 2. Measures relating the voicing period at C1 onset, VTT and C1 voicing and duration in CC#C sequences

	Voicing at C1 onset/C1 voicing ratios, %			
	mean	SD		
p	99.6	1.25		
t	100.0	0.00		
k	99.7	0.87		
s	97.7	3.53		
	Voicing at C1 onset, ms		**VTT, ms**	
	mean	SD	mean	SD
p	18.4	12.48	7.1	4.62
t	24.3	14.14	15.0	11.03
k	20.2	9.76	6.4	2.82
s	12.4	5.68	8.1	5.56
	Voicing at C1 onset/C1 duration ratios, %		**VTT/C1 duration ratios**	
	mean	SD	mean	SD
p	32.4	20.90	11.0	8.51
t	43.1	23.37	10.2	7.25
k	35.2	16.28	21.0	16.19
s	16.1	8.10	7.6	3.24

Data presented in the left half correspond to clusters with a voiced C2 and those presented in the right half to clusters with a voiceless C2.

in fig. 11); also, nasals, lateral and the trill /r/ triggered less regressive voicing than stops and /j/ (not shown).

Another factor which contributes to variations in syllable-final obstruent voicing is the number of consonants in the cluster. A comparison between the frequency of occurrence data plotted in figure 11 and in figure 3 reveals indeed less voicing for word-final obstruents in CC#C sequences than in C#C sequences. As pointed out in section 1.2, the rationale for this difference appears to be that the intraoral pressure level increases with the number of consonants in the cluster. As for the role of consonant position within the cluster and as revealed by figure 11, stops but not fricatives behave as predicted in exhibiting less voicing in C2 than in C1 position (see section 1.2).

The methodology for determining whether, in spite of having low voicing ratios, C1 is planned as voiced in CC#C sequences coincides with the criterion applied to the C#C sequences (see section 3.2.1). First, for instances of partial C1 voicing, the duration of the voicing period at the C1 left edge was compared with VTT duration. Results reveal that voicing at C1 onset accounts for near 100% of the overall C1 voicing duration (table 2, top). The remaining panels of table 2 and the graphs in figure 12 show that voicing at C1 onset is significantly longer than VTT [$F(1, 1{,}314) = 297.97$, $p < 0.0001$ for the corresponding voicing ratios]. Moreover, according to two significant C1 × condition and speaker × condition interactions [$F(3, 1{,}314) = 20.8$, $p < 0.001$; $F(7, 1{,}314) = 8.01$, $p < 0.001$], this difference holds for all speakers

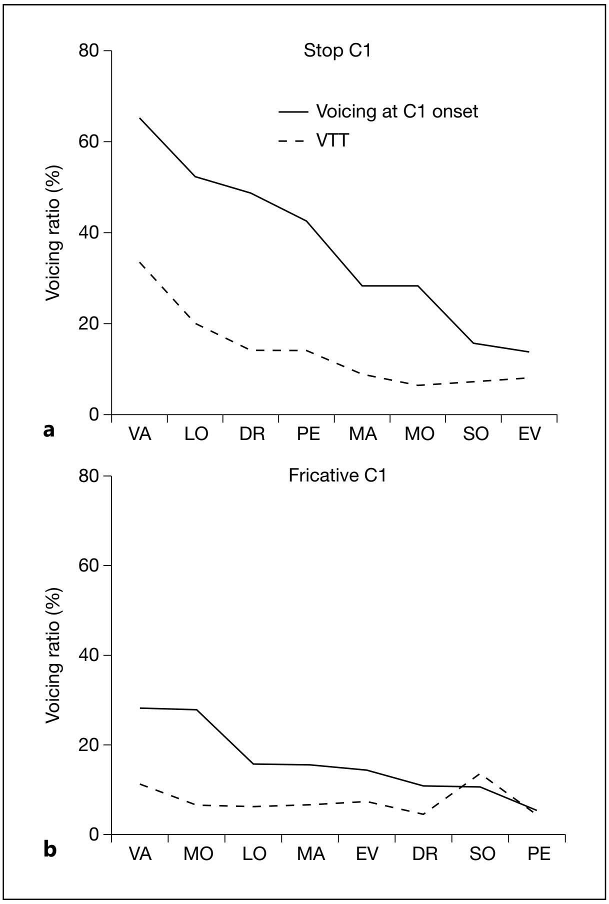

Fig. 12. Speaker-dependent voicing ratios for the voicing period occurring at C1 onset in CC#C sequences compared with the VTT values. Data are plotted separately for clusters beginning with stops (**a**) and fricatives (**b**). In the two graphs, speaker-dependent voicing values at C1 onset are given from highest to lowest.

with the exception of S.O. and is greater for clusters with a stop C1 than for those with a fricative C1 (fig. 12). In the second place, we computed all cluster tokens where C1 may be considered to be planned as voiced by adding the number of repetitions exhibiting 100% voicing to those showing partial voicing and an initial voicing period which was longer than VTT. Assuming that this C1 voicing condition ought to occur at least 80% of the time for voicing assimilation to take place, results plotted in figure 13 may be taken to indicate that assimilation is at work for speakers M.A., L.O. and M.O., but not for subjects P.E. and S.O. According to the data for the remaining subjects, assimilation appears to be transparent for one obstruent but not for the other, i.e., for stops in the case of speakers D.R. and V.A. and for fricatives in the case of E.V.

Statistical results for the segmental duration data reveal that C2 is longer when C3 is voiceless than when it is voiced in the case of all speakers [$F(1, 288) = 27.13$, $p < 0.001$]. As shown in figure 14 (right graphs), this effect occurs generally in stop + fricative + stop combinations but not in the /l/ + stop + fricative sequences. C1 and vowel duration, however, does not change with the C3 voicing status, i.e., C1 is not

DOI: 10.1159/000369028

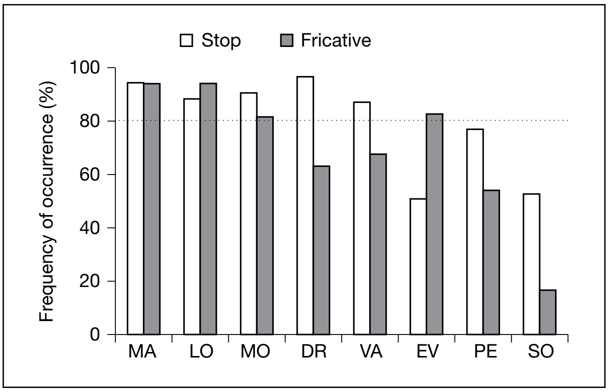

Fig. 13. Frequency of occurrence of estimated regressive assimilation cases in CC#C sequences plotted as a function of C1 manner of articulation and speaker. A no-assimilation baseline has been inserted at the 80% frequency of occurrence.

longer and the vowel is not shorter when C3 is voiceless than when it is voiced (see middle and left graphs). As revealed by a significant speaker × C3 interaction [$F(3, 296) = 4.94$, $p < 0.01$], the only exception appears to be subject M.O., who exhibits a longer vowel in obstruent clusters with a voiced versus voiceless C3.

In sum, voicing in word-final obstruents in Catalan CC#C sequences is less than in C#C sequences and varies as a function of essentially the same target and contextual consonants and speakers. An evaluation of the cluster tokens which are eligible for assimilation reveals a less clear-cut picture for CC#C than for C#C sequences: 3 speakers may be said to assimilate C1 to C3 in voicing, 2 other speakers do not appear to assimilate, and the remaining 3 subjects exhibit C1-to-C3 assimilation, but only clearly so for one of the two consonant manner of articulation categories under analysis. Moreover, C1 and preceding vowel duration seem to play no major role in cueing the voicing assimilation process in three-consonant clusters.

4 Discussion

Experimental findings reported in this article reveal that, for all Catalan speakers subject to analysis, the phonetic implementation of the regressive voicing adaptation process in heterosyllabic consonant clusters is gradient and thus, subject to multiple conditioning factors which cause word-final obstruents to exhibit several voicing degrees. These factors were the manner of articulation characteristics of the target and contextual consonants, and the number of consonants and the consonant position in the cluster. Moreover, speakers differed substantially regarding voicing degree in the target obstruent. In spite of this considerable degree of variability, however, all 8 Catalan subjects under analysis appear to have a regressive voicing assimilation rule in C#C clusters and, though less clearly so, most speakers may be said to exhibit the rule in CC#C sequences as well. For most speakers, whether there was considerable or little voicing at C1 onset, this voicing period was longer than VTT and may thus be said to be associated with the word-initial voicing trigger rather than with the vowel preceding the cluster. Cuartero (2001) also reports longer than VTT voicing periods at the C1 left edge in obstruent#voiced consonant sequences produced by 2 Eastern Catalan speakers independently of whether their consonant voicing ratios were high

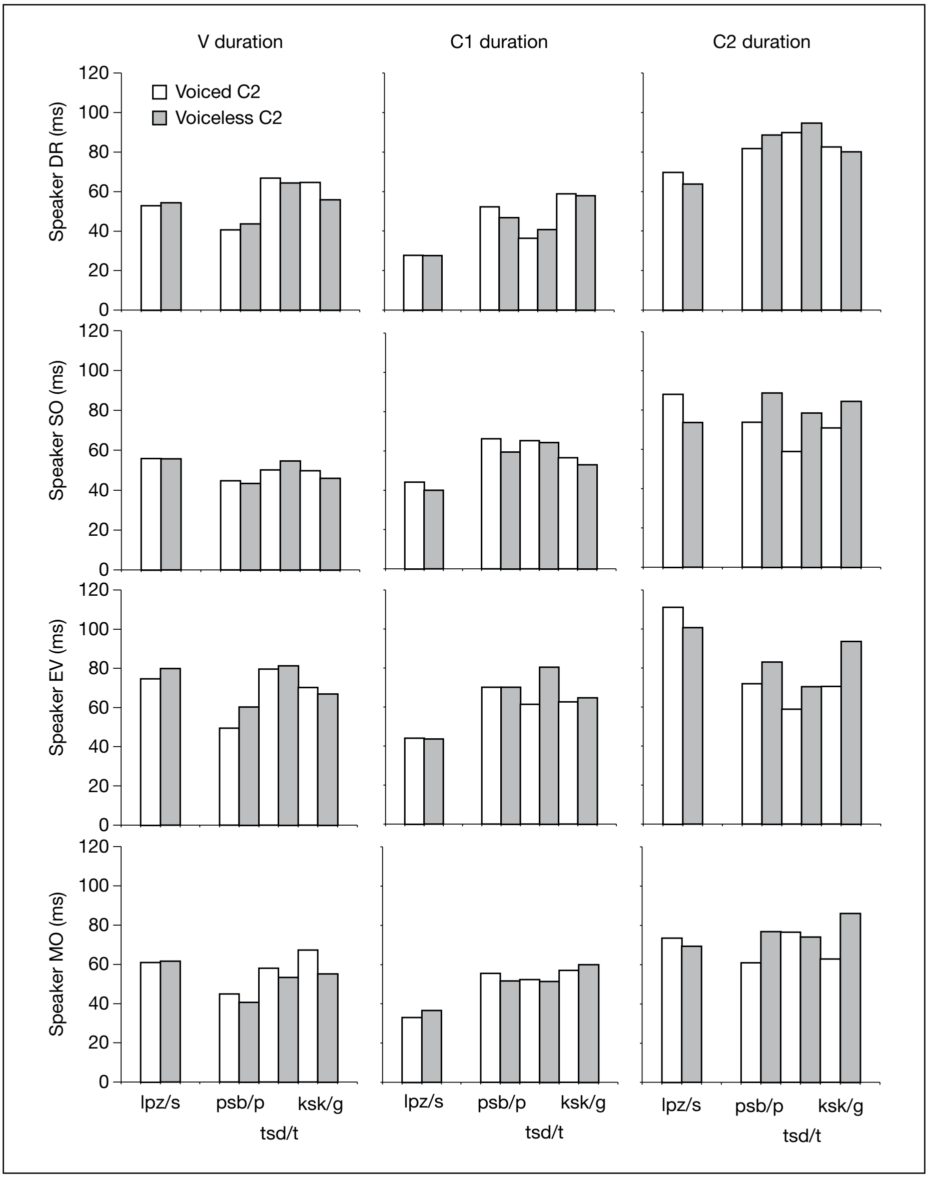

Fig. 14. C2, C1 and preceding vowel duration for /l/ + stop + fricative and stop + fricative + stop cluster pairs for all 4 speakers subject to analysis.

or low. Segmental duration was used by most speakers to cue voicing in word-final obstruents which are adjacent to but not distant from the voicing trigger, i.e., C1 and the preceding vowel in C#C clusters, and C2 but not C1 or the preceding vowel in CC#C sequences. At least 1 speaker, i.e., S.O., exhibited no signs of voicing assimilation in CC#C sequences, which suggests that voicing adaptation for this subject is a

DOI: 10.1159/000369028

short-range coarticulatory process which only operates on the immediately preceding obstruent. Judging from data from other studies it seems that speakers of languages with prevoiced stops like Catalan quite regularly exhibit regressive voicing assimilation. This scenario differs from that for speakers of languages without prevoiced stops such as English where mixed scenarios involving the presence and absence of assimilation are possible. It is also in agreement with articulatory data revealing that the phonetic implementation of regressive place assimilation in /n/ + stop clusters is regularly categorical in some Romance languages (Italian, Spanish) but not in English or German.

In view of the considerable speaker-dependent phonetic variability, the conclusion that practically all 8 Catalan speakers under analysis appear to have a regressive voicing assimilation rule should be accepted with some reservations. Thus, if we had used a different measure of VTT such as VTT + two standard deviations (Slis, 1986), we would have probably concluded that regressive voicing assimilation operates in C#C but not in CC#C sequences at least in the case of speakers showing low C1 voicing degrees. It should be noticed in this respect that, even according to the criteria applied in the present study, the voicing assimilation threshold for C1 in CC#C sequences was not reached by all obstruents for most speakers, and that stops and fricatives occurring in C2 position in these clusters exhibited very low voicing degrees for all subjects. The conclusion that even in these circumstances most Catalan speakers have a voicing assimilation rule seems justified by the fact that obstruent voicing may be cued by different phonetic characteristics acting simultaneously (Lisker, 1986), and also by the belief that a significant increase in voicing at C1 onset with respect to the VTT values (even if the latter are especially short as in the present data set) should be enough to signal C1 as voiced.

There was considerable speaker-dependent and consonant-dependent variability in the phonetic implementation of the voicing assimilatory process. Subjects could be split into at least two groups depending on whether word-final obstruents showed maximal or minimal voicing most of the time, with cases of intermediate voicing adaptation occurring much less frequently. Moreover, speakers who voiced least exhibited more token-to-token variability than those who voiced most and therefore were less consistent in assigning specific voicing degrees to word-final obstruents in clusters. It is hard to ascertain whether such speaker-dependent voicing differences are dialect-dependent or not. This could be so since consonant voicing ratios were less for speakers from certain dialectal areas (P.E. and S.O. from the city of Barcelona, and E.V. and M.A. from Western Catalan) than for those from Eastern Catalan localities other than Barcelona (V.A., D.R., L.O. and M.O.). Data reported in Recasens and Mira (2012) show, however, that these dialect-dependent differences in voicing hold neither for the voice onset time values for voiced stops nor for the degree of voicing for prevocalic word-final fricatives, which are realized as voiced in Catalan. Moreover, speaker-dependent differences in voicing duration at C1 onset in C#C and CC#C sequences did not conform too closely to the corresponding obstruent voicing ratios (compare fig. 8 and 12 with fig. 1, 3–5, 11).

Word-final fricatives showed less voicing and were more variable regarding voicing degree than word-final stops, which appears to be in line with the conflicting requirements involved in combining voicing with frication. They also exhibited less voicing and were more variable than word-initial fricatives probably since voicing is actively controlled for obstruents occurring in the latter versus former position in

Catalan. Data for CC#C sequences in figure 13 indicate that voicing for a fricative C1 may be below the no-assimilatory borderline for 4 out of the 8 speakers under analysis, which may be taken as indicative that constraints on fricative production overcome the speakers' intention to realize C1 as voiced.

Another relevant finding concerns the role that manner of articulation requirements for word-initial voiced consonants play in the regressive voicing adaptation process. Catalan, a language where regressive voicing assimilation is supposed to occur before all sonorants, shows that the degree of voicing in word-final obstruents is less before nasals, laterals and the trill than before approximants, and that fricatives are most reluctant to maintain voicing in these contextual conditions. These data suggest that, among languages with prevoiced stops, the fact that regressive voicing before sonorants operates in some languages but not others may reflect different stages of a sound change process: at its initial stages, the absence of voicing obeys phonetic constraints, i.e., obstruents show less voicing before specific sonorants than before others; at the last stages, these phonetic effects become phonologized and apply to all sonorants, i.e., none of the sonorants available in the language act as voicing triggers. We believe that this approach has some advantages over the proposal that sonorants should not trigger voicing because they are unspecified for voicing underlyingly (Jansen, 2004) in that it is more compatible with the existence of languages like Catalan where, while all sonorants appear to be voicing triggers, voicing degree in word-final obstruents varies depending on the specific manner of articulation characteristics of the following sonorant.

A explained in some detail in section 1.2, the reason why specific sonorants cause little voicing to occur in preceding obstruents could be sought in the degree of compatibility between the articulatory specification for the consonants in succession. Indeed, specific production requirements for nasals, laterals and the alveolar trill prevent much articulatory and voicing anticipation from occurring during a preceding obstruent if its integrity is to be kept. According to this view, obstruents would not be better voicing triggers than sonorants because the pharyngeal cavity may be actively expanded in the former context condition versus the latter (Steriade, 1995), but because of the ways the manner of articulation requirements for the two consonant classes interact with each other. Indeed, frication and an intraoral pressure buildup for stops cannot co-occur easily with nasality, laterality and trilling. This explanation appears to be compatible with data reported in the present investigation showing more token-to-token variability in voicing and segmental duration for word-final obstruents (mostly fricatives) before specific sonorants than before obstruents. These data suggest, for example, that considerable random variability in voicing and segmental duration for fricative + nasal sequences may be related to the variability in velar lowering anticipation during the fricative.

An aspect to be explored in future research is the possibility that languages with prevoiced stops and a regressive voicing assimilation rule differ regarding voicing degree and variability for word-final obstruents in consonant clusters. According to literature reports, voicing assimilation applies more gradiently in Hungarian (also in Catalan judging from the data reported in the present study) than in Russian. It could also be that the extent to which speakers use vocal fold vibration and segmental duration for cueing word-final obstruent voicing in regressive voicing assimilation scenarios varies from language to language.

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. I would also like to thank the reviewers' comments on a previous version of the manuscript.

References

Abadal S, Recasens D (2009): Peakdet2: An Instructions Manual. http://voiceresearch.free.fr/egg.

Abdelli-Beruh NB (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.

Baltazani M (2006): On /s/-voicing in Greek. Proc 7th Int Conf on Greek Linguistics, York, pp 1–12.

Bárkányi Z, Kiss ZG (2014): Phonological categoricality vs. phonetic gradience: the laryngeal properties of Slovak three-consonant clusters. 11th Old World Conf in Phonology, Leiden.

Burton MW, Robblee KE (1997): A phonetic analysis of voicing assimilation in Russian. J Phon 25:97–114.

Celata Ch, Calamai S, Ricci I, Bertini Ch (2013): Nasal place assimilation between phonetics and phonology: an EPG study of Italian nasal-to-velar clusters. J Phon 41:88–100.

Cuartero N (2001): Voicing Assimilation in Catalan and English; PhD diss Universitat Autònoma de Barcelona, Barcelona.

Docherty GJ (1992): The Timing of Voicing in British English Obstruents. Dordrecht, Foris.

Dvořák V (2010): Voicing assimilation in Czech. Rutgers Working Papers in Linguist 3:115–144.

Ellis L, Hardcastle WJ (2002): Categorical and gradient properties of assimilation in alveolar to velar sequences: evidence from EPG and EMA data. J Phon 30:373–396.

Farnetani E, Busà MG (1994): Italian clusters in continuous speech. Proc 1994 Int Conf on Spoken Language Processing, Yokohama, pp 359–362.

Gow DW, Im AM (2004): A cross-linguistic examination of assimilation context effects. J Mem Lang 51:279–296.

Hallé PA, Adda-Decker M (2011): Voice assimilation in French obstruents: categorical or gradient? In Goldsmith J, Hume E, Wetzels L (eds): Tones and Features. Phonetic and Phonological Perspectives. Berlin, de Gruyter, pp 149–175.

Jansen W (2004): Laryngeal Contrast and Phonetic Voicing: A Laboratory Phonology Approach to English, Hungarian, and Dutch; doct diss Rijksuniversiteit Groningen.

Jansen W (2007a): Dutch regressive voicing assimilation as a 'low level phonetic process: acoustic evidence; in van de Weijer J, van der Torre EJ (eds): Voicing in Dutch: (De)voicing – Phonology, Phonetics, and Psycholinguistics. Amsterdam, Benjamins, pp 125–152.

Jansen W (2007b): Phonological 'voicing', phonetic voicing and assimilation in English. Lang Sci 29:270–293.

Jansen W, Toft Z (2002): On sounds that like to be paired (after all): an acoustic investigation of Hungarian voicing assimilation. SOAS Working Papers in Linguist 12:19–52.

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

Kulikov V (2013): Voicing contrast in consonant clusters: evidence against 'sonorant transparency to voice assimilation' in Russian. Phonology 30:423–452.

Lisker L (1986): 'Voicing' in English: a catalogue of acoustic features signaling /b/ versus /p/ in trochees. Lang Speech 29:3–11.

Markó A, Gráczi TE, Bóna J (2010): The realisation of voicing assimilation rules in Hungarian spontaneous and read speech: case studies. Acta Ling Hung 57:210–238.

Ohala JJ, Solé MJ (2010): Turbulence and phonology; in Fuchs S, Toda M, Zygis M (eds): Turbulent Sounds: An Interdisciplinary Guide. Berlin, de Gruyter, pp 37–97.

Recasens D, Mira M (2012): Voicing assimilation in Catalan two-consonant clusters. J Phon 40:639–654.

Recasens D, Mira M (2013): Voicing assimilation in Catalan three-consonant clusters. J Phon 41:264–280.

Rothenberg M, Mahshie JJ (1988): Monitoring vocal fold abduction through vocal fold contact area. J Speech Hear Res 312:338–351.

Schmidt LB, Willis EW (2011): Systematic investigation of voicing assimilation of Spanish /s/ in Mexico City; in Alvord SM (ed): Selected Proc of the 5th Conf on Laboratory Approaches to Romance Phonology, Somerville, Cascadilla Proceedings Project, pp 1–20.

Slis IH (1986): Assimilation of voice in Dutch as a function of stress, word boundaries and sex of speaker and listener. J Phon 14:311–326.

Smith CL (1997): The devoicing of /z/ in American English: effects of local and prosodic context. J Phon 25:471–500.

Solé MJ (2002): Aerodynamic characteristics of trills and phonological patterning. J Phon 30:655–688.

Solé MJ (2007): The stability of phonological features within and across segments: the effect of nasalization on frication; in Prieto P, Mascaró J, Solé MJ (eds): Segmental and Prosodic Issues in Romance Phonology. Amsterdam, Benjamins, pp 41–66.

Solé MJ (2009): Assimilatory processes and aerodynamic factors in the interaction of features: the case of nasality and voicing; in Vigário M, Frota S, Freitas MJ (eds): Phonetics and Phonology: Interactions and Interrelations. Amsterdam, Benjamins, pp 205–234.
Steriade D (1995): Underspecification and markedness; in Goldsmith JA (ed): Handbook of Phonological Theory. Oxford, Blackwell, pp 114–174.
Strycharczuk P (2012): Phonetics–phonology interactions in pre-sonorant voicing; doct diss University of Manchester, Manchester.
Westbury J (1979): Aspects of the temporal control of voicing in consonants in English. Texas Linguistic Forum, 14. Austin, Department of Linguistics, University of Texas.
Westbury J, Keating P (1986): On the naturalness of consonant voicing. J Linguist 22:145–166.
Wetzels WL, Mascaró J (2001): The typology of voicing and devoicing. Language 77:207–244.
Wheeler M (2005): The Phonology of Catalan. Oxford, Oxford University Press.