Author's personal copy

Journal of Phonetics 42 (2014) 37–51

Contents lists available at ScienceDirect

Journal of Phonetics

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

Acoustic characteristics of (alveolo)palatal stop consonants, and velar softening

Daniel Recasens [a,b,*]

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

A R T I C L E I N F O

Article history:
Received 12 November 2012
Received in revised form
4 December 2013
Accepted 12 December 2013
Available online 15 January 2014

A B S T R A C T

The paper investigates using data from Majorcan Catalan the acoustic characteristics, and the vowel context and positional conditions, that contribute to the identification of the unaspirated (alveolo)palatal allophone [c] of /k/ as the palatoalveolar affricate /tʃ/ by listeners, and therefore to the implementation of velar softening in the world's languages. Results from perception tests run on [cV] excerpts reveal that affricate percepts are more likely to occur when the (alveolo)palatal stop appears before /i/ than before /a/, which is in agreement with universal patterns of velar softening, and in word-initial and word-final intervocalic position than word-medial intervocalic and utterance initially. Utterance finally [c] is prone to be heard as the fricative [ç]. Affricate identification appears to be associated with context- and position-dependent acoustic cues: high frequency F2 vowel transition endpoints and stop burst spectra, and a long burst, before /i/ and word initially; long range F2 vowel transitions next to /a/, and an intense stop burst in this same vowel context and in intervocalic position. High /tʃ/ identification percentages for [c] in the sequence [uc#u], as well as differences among speakers in producing affricate-like realizations of [c] and among listeners in perceiving the stop as an affricate, are also discussed.

© 2013 Elsevier Ltd. All rights reserved.

1. Introduction

This paper is about the phonetic motivation of velar softening, namely, a sound change involving the shift of /k/ into the affricate /tʃ/ in specific contextual and positional conditions (Italian [ˈtʃɛnto]<Latin /ˈkɛnto/ "one hundred"). The main goal of this investigation is to show that the perceptual identification of an unaspirated velar stop as /tʃ/ may be accounted for through the intermediate (alveolo)palatal stop realization [c], and to uncover the acoustic characteristics which contribute to velar softening. In principle, this sound change process may occur in languages or dialects where [c] is an allophone of the velar stop phoneme /k/, e.g., in Majorcan Catalan (a dialect of Catalan spoken in the Balearic island of Majorca which will be subjected to analysis in the present investigation), but also in Greek dialects (Syrika, Kong, & Edwards, 2011), and in Romance dialectal domains such as Picard ([tʃjɛr]<[cjɛr] Latin /ˈkaro/ "expensive"; Remacle, 1953: map 16), Romansh and Northern Italian (see examples below). The change of phonemic /c/ into an affricate may be exemplified by Czech dialects where the stop release has been characterized as clearly affricated (Šimáčková, Podlipský, & Chládková, 2012). It should be noticed at this point that the occurrence of the consonant sound [c] in the world's languages is not as infrequent as it may seem: according to the P-base (Mielke, 2007), it appears in 100 languages, more frequently than the lateral cognate [ʎ] (47) and less often than the (alveolo)palatal nasal [ɲ] (270) and the velar stop [k] (580). Also according to Maddieson (1984), the frequency of occurrence of these four consonants decreases in the progression [k] (283 languages)>[ɲ] (107)>[c] (41)>[ʎ] (15).

Depending on consonant manner of articulation, dialect and speaker, (alveolo)palatal consonants such as [c] may be articulated at the alveolopalatal zone with the tongue blade and predorsum, or else solely at the palatal zone with the dorsum of the tongue. In the former case, closure or constriction takes place simultaneously at the alveolar and palatal areas, and in the latter at the prepalato-mediopalatal or medio-postpalatal areas (Recasens, 2013). Any of these realizations of the (alveolo)palatal oral stop differs crucially from the front allophone of /k/ before a front vowel or glide ([k₊]) in that the latter is postpalato-velar and therefore articulated at the rear of the hard palate exclusively (see Corneau, Soquet, & Demolin, 2000 for French, and Recasens & Pallarès, 2001 for Eastern Catalan). This means that, whenever evolving from front /k/, the stop must change place of articulation from velar to (alveolo)palatal in order to become [c]. At the acoustic level, differences between front /k/ and [c] are to be sought mainly in the stop burst (see Section 1.1, and also data for Czech, Hungarian and Eastern Catalan in Blumstein, 1986, Keating & Lahiri, 1993, and Zygis,

* Correspondence address: Departament de Filologia Catalana, Universitat Autònoma de Barcelona, Bellaterra, Barcelona.
E-mail address: daniel.recasens@uab.es

0095-4470/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.wocn.2013.12.003

Author's personal copy

Recasens, & Espinosa, 2010): a more anterior closure location for [c] than for front /k/ causes the front-cavity dependent burst spectral peak to rise typically from about 3000 Hz to about 4000 Hz; the stop burst is usually longer for [c] (up to about 60 ms or even higher) than for front /k/ (below 40–50 ms) in line with differences in primary articulator and dorsopalatal contact size between the two consonant realizations.

There are reasons to believe that intermediate (alveolo)palatal stop realizations are required for the velar stop to be confused with /tʃ/. Firstly, perception data summarized below reveal that [c] may be confused with /tʃ/ whenever the stop burst is sufficiently salient (Recasens & Espinosa, 2009). Secondly, velar softening has operated not only before a front vocalic segment but also before /a/ and word finally where [c] does not clearly derive from front /k/. While the former context is the most common velar softening trigger in the world's languages (Romance, Slavic, Indo-Iranian, Bantu; Guion, 1998; Bhat, 1978), the two latter ones may also cause velar softening to apply (Picard [cjɛr, tʃjɛr] and French *cher* [ʃɛʀ] from Latin /ˈkaro/ CARU “expensive”, Sutselvan Romansh [lac, latʃ] from Latin LACTE “milk”; Pope, 1934; Luzi, 1904: 811).

Finally, the stop and the affricate may co-occur in specific Romance dialectal areas, as exemplified by the Picard and Romansh phonetic variants mentioned below and other ones like Fassan Ladin [ˈcawra]/[ˈtʃawra] derived from Latin /ˈkapra/ CAPRA “goat” (Elwert, 1943: 67).

This hypothesis is in contrast with the view that the perceptual replacement of front /k/ by /tʃ/ cannot be triggered by a change in place of articulation since closure location and the primary articulator for the velar stop and for the affricate are discontinuous (Ohala, 1992). Following this line of thought, it was proposed that velar softening is achieved through perceptual confusion between the two sounds given the similarity between the 2500–3500 Hz spectral peak for both the front velar stop burst and the affricate frication noise (Guion, 1996, 1998). In support of this rationale, Guion (1998) found that phonetic realizations of aspirated front /k/ may be confused with /tʃ/ by listeners mostly when the duration of the burst frication and aspiration period is increased significantly. A possible problem with Guion's experimental setup is that these perceptual confusions occurred only when a high amplitude white noise was superimposed on the perception stimuli which casts some doubts as to whether these stimuli were natural enough. Moreover, analysis and perception data reported in Zygis et al. (2010) show that the burst frication period for the front allophone of German aspirated /k/ is not sufficiently long and intense to be integrated as an affricate both by German listeners and by informants of Catalan where /k/ is unaspirated. A related acoustic-based explanation of velar softening has been proposed: whenever the burst frequency peak for front /k/ is degraded, listeners may hear /t/ instead of /k/, and even /tʃ/ rather than /t/ if the stop is strongly aspirated (Chang, Plauché, & Ohala, 2001). As argued above, the present paper holds the alternative view that the change /k/>/tʃ/ is mediated by the articulatory realization [c] which may share the same closure location, primary articulator and spectral characteristics as the affricate.

As analyzed in detail later on, the path from [c] to /tʃ/ may be sought, at least in part, in what Garrett and Johnson (2010) call ‘phonetic enhancement effect’ by which the intensification of a specific acoustic characteristic through some articulatory and aerodynamic action may contribute to sound change implementation. Along these lines, we hypothesize that an increase in constriction degree and airflow volume in specific contextual and positional conditions may cause an increase in the spectral frequency of the (alveolo)palatal stop burst and the vowel transition endpoints, and in burst duration and intensity, thus contributing to the categorization of [c] as an affricate.

Within this theoretical framework, the main goal of this investigation is to determine by means of acoustic analysis and perceptual evaluation the acoustic cues that cause the unaspirated (alveolo)palatal stop [c] to be identified as /tʃ/, as well as those contextual vocalic and positional conditions which favor the implementation of velar softening. This research will be carried out on acoustic data for Majorcan Catalan where the allophone [c] of /k/ occurs in a large number of vowel contexts and positions. In this Catalan dialect, /k/ is realized as [k] before a back rounded vowel word internally ([kus ˈfɔrsə] cus força “he/she sows a lot”), and as [c] before /i, e, ɛ, a, ə, j/ ([ci ˈve] qui ve “who is coming?”, [sə ˈcazə] sa casa “the house”, [cəˈva] cavar “to dig”), and word finally before a pause or the initial vowel of the following word ([bɔn ˈsuc] bon suc “good juice”, [saˈc amplə] sac ample “wide sack”) (Veny, 1978). This situation is in contrast with the one occurring in Eastern Catalan (the most common Catalan dialect) where, analogously to French or English, /k/ is realized through a single allophone which may be slightly more anterior before front vowels (palatovelar) than before low and back vowels (velar) (Recasens & Pallarès, 2001). Moreover, both in Eastern Catalan and Majorcan, the affricate /tʃ/ occurs in the word-medial intervocalic and word final positions ([ˈdutʃə] dutxa “shower”, [ˈfitʃə] fitxa “card, counter”, [bɔtʃ] boig “mad, masc.”, [matʃ] maig “May”), as well as word initially and postconsonantally in alternation with [ʃ] depending on lexical item, dialectal area and speaker ([(t)ʃukuˈlatə] xocolata “chocolate”, [punˈ(t)ʃa] punxar “to poke, to string”).

1.1. Recasens and Espinosa (2009)

This paper is a continuation of Recasens and Espinosa (2009) where analysis data and affricate identification results were reported for productions of Majorcan Catalan [c]. In this study, production data for the (alveolo)palatal stop were analyzed utterance initially and utterance finally next to /i, a, u/, and in symmetrical V#CV sequences with /i, a/ and in symmetrical VC#V sequences with /u/. Identification results were obtained from a forced choice perceptual test where Majorcan and Eastern Catalan informants were asked to identify [cV] and [Vc] excerpts excised from the productions mentioned above as either /k/ or /tʃ/. Affricate responses for the [c] stimuli were consistent with the following common spectral characteristics between the two sets of sounds referred to next (data for /tʃ/ in Majorcan and Eastern Catalan reported in the present study have been taken from Recasens & Espinosa (2009) and Recasens (1986)):

(a) A front-cavity burst spectral peak for [c] at about 2500–4000 Hz before front and low vowels and about 2000–3500 Hz before back rounded vowels, which is close to the 3000–4000 Hz front cavity-dependent spectral peak of the affricate frication noise.
(b) F2 vowel transition endpoint frequencies in the context of /i/ (2200–2500 Hz), /a/ (1700–2000 Hz) and /u/ (1100–1500 Hz), which match those for the affricate (1500–2000 Hz) at least in the case of the two former vowel contexts, and are positively related to tongue dorsum raising and dorsopalatal contact size and inversely related to lip rounding (Fant, 1960).

It should be mentioned at this stage that, when describing the similarities between [c] and /tʃ/, we are mainly interested in the context- and position-dependent variations that may cause [c] to approach /tʃ/, but less so in the contextual changes that may occur during /tʃ/. This methodological criterion is based on the assumption that, as long as the stop approaches the affricate borderline regarding a specific phonetic property, listeners may be prone to identify the former as the latter.

The frequency characteristics referred to above were found to be necessary but not sufficient cues for affricate identification. Indeed, while [c] and [tʃ] share similar spectral properties, an increase in burst duration and intensity must take place for the (alveolo)palatal realization of /k/ to be confused with /tʃ/: while the unaspirated stop burst is usually less than 40 ms long and its energy level relative to that for the adjacent vowel is about 0.65–0.75 dB, the affricate frication period lasts for about 70 ms or longer and its relative energy amounts to 0.75–0.95 dB. The effectiveness of these

Author's personal copy

acoustic properties explains why the affricate identification percentages in our previous study were higher in some context and position conditions than in others:

(a) Utterance finally, presumably due to final lengthening which affects prepausal vowels and consonants in Catalan (Recasens, 1991). Indeed, the burst was found to be at least 70–80 ms long, and thus approached the duration of the affricate frication noise, in utterance final position as opposed to less than 40 ms in other positions.
(b) Intervocalically in the sequences /a#ka, uk#u/, and utterance initially before /a/ (/#ka/) as well. Affricate identification in this case should be attributed mostly to burst energy which turned out to be higher intervocalically than utterance initially and utterance finally (also Zygis et al., 2010), and in the context of the more open vowel /a/ and, less so, /u/ than in the context of /i/. The articulatory factors causing these context- and position-dependent differences in burst intensity appear to be front cavity size which is greater for low than for front vowels, and the cross-sectional area of the lingual constriction which should be especially large in intervocalic position (Dorman, Studdert-Kennedy, & Raphael, 1977; Stevens, 1998). In addition to burst intensity, the presence of large F2 vowel transition ranges, i.e., the formant frequency difference between the transition endpoint and the vowel steady-state period, should also contribute to an increase in affricate identification percentages in sequences with /a, u/. The active role of the vowel transitions following the stop closure in affricate identification in the low vowel context appears to be consistent with [c] exhibiting an [j]-like configuration and a high F2 frequency transition endpoint at consonant offset, as well as with kinematic data for the sequence [aca] showing that the CV lingual movement trajectories are longer and exhibit lesser displacement and velocity than the VC trajectories (Recasens & Espinosa, 2009, 2010b). Co-occurring frication during these long-range CV vowel transitions when the stop constriction becomes especially narrow could very well cause the [j]-like element to be integrated as the frication phase of an affricate.

As pointed out in detail in the following Sections 1.1.1–1.1.3, there are good reasons for conducting a new study that improves significantly the design and results of the analysis and perception experiments performed in Recasens & Espinosa (2009): while [c] yielded a considerable number of affricate responses in our previous study, those responses were not associated with the front vowel context, which favors velar softening in the world's languages (Section 1.1.1); the positional conditions under which affricate identification occurs remained unclear partly because [c] was recorded in a small subset of word and utterance positions only (Section 1.1.2); some methodological steps which were taken in the previous investigation may have rendered the speech material subjected to analysis and perceptual evaluation unnatural in some respects (Section 1.1.3).

1.1.1. Vowel context

Our previous study failed to replicate the universal trend for velar softening to operate before front vocalic segments rather than before low and back ones. The poor affricate identification results obtained for [c] before /i/ were attributed to the presence of a short and weak stop burst, which implies that none of the Majorcan Catalan speakers whose data were used for the preparation of the perceptual stimuli produced the (alveolo)palatal stop before /i/ with enough constriction narrowing and frication so that it could sound like an affricate. This negative outcome calls for the replication of the study with data for a different set of speakers. Therefore, a first hypothesis of the present investigation is that the categorization of [c] as an affricate should be most prone to occur before /i/, and that this ought to be so to a large extent since the [c] burst is expected to be longer before this vowel than before other vowels (regarding an increase in burst duration for lingual stops before high vs low vowels, see also Docherty, 1992). An increase in burst duration before /i/ should be related to the formation of a long and narrow constriction for the passage of airflow at consonant release when the stop consonant is articulated with a considerable degree of linguopalatal contact. Moreover, in line with the fact that the affricate frication noise may exhibit a higher frequency peak than the frication period of [c] (see above), a higher burst spectral frequency and a higher F2 at the endpoint of the vowel transitions for the (alveolo)palatal oral stop before /i/ than before /a, u/ could also contribute to velar softening in the former context rather than in the two latter ones.

Affricate categorization ought to take place before /a/ as well (though less so than before /i/). This expected outcome is in accordance with sound change data from the Romance languages showing that velar softening may operate in this vowel context, and with analysis data revealing the presence of several acoustic characteristics which could render the stop affricate-like in the low vowel context condition (see Section 1.1): large F2 vowel transition ranges, which may contribute to the integration of [j] as the fricative element of an affricate if the lingual constriction is sufficiently narrow; a high stop burst energy level; relatively high frequency burst spectra and F2 vowel transition endpoints.

1.1.2. Position

Another reason for carrying out a new study is to explore more thoroughly the hierarchy of positions under which affrication occurs. As summarized above, affricate responses in our previous study turned out to be as high or higher utterance finally and intervocalically (both in the case of the V#CV and VC#V sequences) than utterance initially, presumably since the stop burst was especially long prepausally and especially intense intervocalically. Assuming that the change of a stop into an affricate involves some articulatory reinforcement (see Buiza & Plug, 2012; Lavoie, 2001 in this respect), these data appear to be in contradiction with the prediction that articulatory strengthening ought to vary with position in the progression utterance initial > word initial > syllable initial, as revealed by literature studies (Fougeron & Keating, 1997) and by electropalatographic data showing that closure duration and linguopalatal contact size for [c] in Majorcan Catalan are greater after a pause than intervocalically (Recasens & Espinosa, 2006, 2009). A rationale for this apparent contradiction follows.

As for the utterance-final scenario (and as confirmed by the auditory impression of the paper author), prepausal [c] yielded high affricate identification percentages presumably since a very long burst caused the stop to be heard as [ç], and this fricative sound was judged to be more similar to /tʃ/ than to /k/ by the informants who took the identification test. This possibility is consistent with word-final [c] becoming [ç] as well as [tʃ] in Romance dialects (Nonsberg [laç] from Latin /ˈlako/ LACU "lake", Romansh [lac, latʃ] from Latin /ˈlakte/ LACTE "milk").

As for the presence of high /tʃ/ identification percentages for intervocalic [c], we would like to propose that the best environment for the identification of [c] as an affricate is the word-initial intervocalic position. The rationale for this proposal is that, while the word initial position contributes to articulatory strengthening through closure and burst lengthening and a frequency increase at the endpoint of the F2 vowel transitions, the presence of two flanking vowels may cause the stop burst to approach the intensity level of /tʃ/. It may thus be argued that an increase in all these acoustic dimensions is needed for [c] to be integrated as an affricate. Moreover, to the extent that phonetic segments may be strengthened word finally (Barnes, 2006; Keating, Wright, & Zhang, 1999), it is also hypothesized that the identification of [c] as an affricate could be favored in word-final intervocalic position as well. The utterance initial and the word-medial intervocalic positions appear to be less likely affrication candidates presumably since the burst energy level may not be high enough in the former case and the closure and frication periods are especially short in the latter. Treating the word-

Author's personal copy

medial intervocalic position as weaker than the word-initial position is consistent with the trend for the velar softening outcome to be an affricate in the latter position and both an affricate and a fricative in the former (Romansh [tʃun] from Latin /ˈkinkwe/ CINQUE "five", [viˈʒin] from Latin /viˈkino/ VICINU "neighbour", Southern Italian dialects [ˈdɛtʃi] from Latin /ˈdɛke/ DECE; Rohlfs, 1966: 289, Lausberg, 1970: 317, 363).

Based on this working hypothesis and on the positional distribution of the affricate and fricative outcomes of original [c] in several Romance languages, the evolutionary path from /k/ to /tʃ/ through [c] could be reconstructed as follows:

(a) The affricate outcomes that we observe word initially and word finally (Romansh [tʃun], [latʃ]) should be generated though the change [c] > [tʃ] in the V#CV and VC#V positional conditions rather than postpausally or prepausally.
(b) The fricative outcome that we observe in the word-medial intervocalic position (Romansh [viˈʒin]) arose either through stop lenition (i.e., [c] > [ʃ]) or through stop affrication followed by lenition (i.e., [c] > [tʃ] > [ʃ]). On the other hand, the fricative outcome that we find word finally (Nonsberg [laç]) may have developed through the straight replacement of [c] by [ç] when the stop was followed by a pause.

The hypothesis that velar softening should vary with position in the ways outlined above could not be fully explored in our previous study since data for the V#CV and VC#V sequences were available only for a subset of vowels, and data for the VCV condition were not included. The present investigation will test the effect of position and vowel context on velar softening with acoustic data for all sequences where the allophone [c] of /k/ occurs in Majorcan Catalan, i.e., before /i, a/ in V#CV, VCV and #CV sequences and after /i, a, u/ word finally whether before a vowel or a pause.

1.1.3. Experimental conditions

The need to perform a new study is also motivated by several methodological issues. Artificial palates were used for the speech recordings in Recasens and Espinosa (2009) which questions the naturalness of the speech material subjected to analysis mostly since the palates in question may have interfered with the quality of the [c] frication noise. Moreover, due to limitations of the EPG data acquisition system, the acoustic signal was digitized at 10 kHz, which did not allow taking spectral measures of the [c] burst above 5000 Hz. Finally, given that the stop burst productions of the Majorcan Catalan speakers who took part in the study were not salient enough, the burst energy level of the acoustic stimuli which were presented to Catalan subjects for perceptual identification had to be increased by about 12 dB so that the stimuli in question could become more affricate-like. The present investigation will cope with these methodological shortcomings by recording natural unperturbed speech at a 22,050 Hz sampling rate, and by keeping the original stop intensity level of the perception stimuli largely unmodified.

1.2. Summary

This present study will look for the acoustic characteristics which render [c] similar to /tʃ/ in specific contexts and positions, and will submit them to perceptual evaluation. It improves in many respects the experimental design of our previous study Recasens and Espinosa (2009) by using a more complete set of contextual and positional conditions, a different set of Majorcan Catalan speakers, and more natural experimental conditions. Two major hypotheses will be tested:

Hypothesis 1. The categorization of [c] as an affricate ought to occur before /i/ in line with the universal pattern of velar softening, and with experimental evidence showing that the stop burst may be especially long and that the burst and the vowel transitions may exhibit high frequency characteristics in this vowel context. To a lesser extent, affricate identification is also expected to take place before /a/ mostly due to the high stop burst intensity level and the large vowel transition ranges.

Hypothesis 2. The identification of [c] as an affricate is expected to occur in the word-initial intervocalic position and, to a lesser extent, in the word-final intervocalic position rather than utterance initially and in the word-medial intervocalic position. This position ranking is based on position-dependent differences in articulatory strengthening and burst intensity.

The paper is divided into two separate sections. The analysis in Section 2 explores the potential contribution of the stop burst spectral, duration and intensity properties, and of the F2 frequency endpoints and ranges of the vowel transitions, in rendering [c] affricate-like in different contexts and positions. Section 3 provides results from two identification tests where the perceptual contribution of these acoustic characteristics to affricate categorization is evaluated.

2. Production

2.1. Method

2.1.1. Recording procedure

Acoustic data were collected for /k/ next to the vowels /i, a, u/ in symmetrical V#CV, VC#V and VCV sequences and in #CV and VC# sequences (see 1–15 in Appendix A). Throughout the paper the first and second vowels in the intervocalic sequences will be often referred to as 'V1' and 'V2', and the five sequence structures V#CV, VC#V, VCV, #CV and VC# as 'sequence types'. The velar stop phoneme exhibits an (alveolo)palatal realization in all cases except for the V#CV, VCV and #CV sequences with postconsonantal /u/ where it is realized as [k] (see Section 1 and phonetic transcriptions in Appendix A).

All segmental sequences subjected to analysis were embedded in three- to five-syllable long Catalan sentences. Lexical stress fell on all vowels adjacent to the target stop consonant except for the VCV sequences 7–9 where V2 but not V1 bears lexical stress since words are required to exhibit only one lexical stress in Catalan. In this case the quality of unstressed V1 differs from that of stressed V2 as follows: /a/ is realized as [a] when stressed (V2) but as [ə] when unstressed (V1); moreover, among the two back rounded vowels [o] and [u] which may occur in unstressed position in Majorcan Catalan, [o] had to be chosen as V1 since meaningful two-syllable words with V1 = [u] were hard to find.

As shown by the speech material listed in Appendix A, the consonants preceding and/or following the V#CV, VC#V, VCV, #CV and VC# sequences are produced for the most part with a relatively high and front tongue body position, i.e., the dentals /t, d/ and the alveolars /s, z, n/. In a few cases and due to lexical restrictions, however, the contextual consonants of interest are articulated with some tongue predorsum lowering (the alveolar trill /r/ in sentences 7 and 13, and dark /l/ in sentence 10), or with lip closing and no prominent tongue activity (the labial /b/ in sentence 9). Differences in vocal tract configuration among these contextual consonants were not expected to affect significantly the frequency characteristics of

Author's personal copy

distant acoustic events subjected to measurement, i.e., the burst spectral peak at /k/ offset and the endpoint of the formant transitions at V1 offset and V2 onset.

The speech material was read several times by 9 middle-aged male speakers born and living in the town of Manacor located in the island of Majorca whose native tongue is Majorcan Catalan and use it exclusively in their everyday life. These subjects were preselected as genuine by the coordinator of the language services department of the Manacor town hall since they show all other most relevant dialectal features and have been living all their lives in the island. The need to carry out a preliminary selection of informants is justified by the fact that the affricate-like quality of [c] in present-day Majorcan Catalan is a non-prestigious speech feature (regarding this and other receding dialectal features such as /l/ darkness and the attitude of Majorcan Catalan speakers towards them, see Simonet, 2010 and Pieras, 1999). Data for those three speakers who produced [c] with a high palatal and thus affricate-like quality, i.e., SI (Sion), LLO (Llodrà) and AN (Andreu), were selected for the present experiment. It should be stated in this respect that the goal of the present investigation was not to characterize the (alveolo)palatal stop productions of the Majorcan Catalan speakers as a whole, but to elicit those acoustic features which may contribute to the identification of [c] as /tʃ/ in those contexts and positions where velar softening occurs in the world's languages. This explains why only those speakers whose productions could trigger the change in question were selected for analysis since, as already found in our previous study, this was prone to occur only whenever specific affrication cues were available in the appropriate contextual and positional conditions. This selection procedure was carried out both though listening and based on acoustic measures mosty of /k/ before /i/ in the sequences V#CV, VC#V, VCV and #CV where velar softening is most prone to occur. The most noticeable speaker-dependent difference appeared to be in burst duration: the burst was especially long (about 55–70 ms; see also Section 3.1) word initially for speakers AN and LLO and utterance initially for speaker SI. The five speakers showing the longest burst of all nine subjects exhibited similar burst spectral characteristics both regarding the front-cavity dependent spectral peak and the center of gravity (the latter measure yielded 5700–5900 Hz in all cases; see Section 2.1.2).

Acoustic recordings were made with a 22,050 Hz sampling rate in quiet recording conditions in the Majorcan town of Manacor using a laptop with a 24-bit external Sound Blaster card and an AKG D70 microphone. Although some higher spectral information may be missed using a 22,050 Hz sampling rate, this sampling frequency ought to be sufficient for capturing all the essential spectral aspects of the stop productions whether derived from single frequency peaks or from the centroid. Five repetitions of each sequence for each speaker were chosen for analysis.

2.1.2. Data analysis

In order to investigate the effect of vowel context and position on consonant realization, several acoustic characteristics were subjected to measurement using the Computerized Speech Lab (CSL) analysis system of Kay Elemetrics, and applying essentially the same criteria as in our previous study Recasens and Espinosa (2009).

The F1, F2 and F3 frequencies were measured at the onset, midpoint and offset of V1 and V2 in the V#CV, VC#V and VCV sequences, and at the vowel endpoint and midpoint in the #CV and VC# sequences. All frequency measurements were taken on spectrographic displays placing a cursor in the middle of the formant with the assistance of LPC spectra whenever necessary. The vowel formant frequency ranges were determined by subtracting the frequency value at the vowel midpoint from that at the vowel edges limiting with the stop consonant. Only data for the F2 vowel transition endpoints and ranges will be submitted to statistical analysis since they reflect most accurately variations in tongue dorsum raising and fronting (Fant, 1960), and are most effective in consonant identification (Liberman, Delattre, Cooper, & Gerstman, 1954; Delattre, Liberman, & Cooper, 1955). Transition ranges may be positive or negative depending on whether the formant frequency at the vowel transition endpoint is higher or lower than that at the vowel midpoint, respectively.

The spectral peak and center of gravity (COG) of the burst for the allophones [c] and [k] of /k/ were measured on autocorrelation 0–11025 Hz LPC spectra with 14 coefficients. Spectra were sampled at the onset of the frication noise following the burst spike with a narrow 15 ms full-Hamming window so as to exclude information located outside the frication period. The burst spectral peak was identified on spectral averages across repetitions of each sequence at about the expected frequencies referred to in Section 1.1. COG was calculated by adding the product of all spectral frequencies by their corresponding amplitudes and dividing the outcoming value by the sum of all the amplitude values (McDonough & Wood, 2008; Ladefoged, 2003). COG values were collected for each sequence token and averaged across all tokens. Both spectral peak and COG are front-cavity dependent and, therefore, should capture similar sources of information (Cho, Jun, & Ladefoged, 2002).

Closure duration for the allophones [c] and [k] of /k/ was measured from the offset of formant structure for the preceding vowel until the stop burst in the V#CV, VC#V, VCV and VC# sequences, and burst duration from the burst spike until the first glottal pulse of the following vowel in the V#CV, VC#V, VCV and #CV sequences and until burst frication offset in the VC# sequences. Absolute burst energy values were computed on 10 ms window energy profiles at two points in time for each sequence token: at the same temporal frame selected for measuring the spectral peak frequency (i.e., immediately after the burst spike) and at burst midpoint, in order to account for the fact that the intensity level may change during the burst. Energy values in dB are obtained by multiplying intensity by duration (Dorman et al., 1977). Relative burst energy values were also calculated for all sequence repetitions by dividing the absolute energy value at the stop burst by that at the midpoint of the stressed vowel following the consonant in V#CV, VC#V, VCV and #CV sequences and preceding the consonant in VC# sequences (Cho et al., 2002).

2.1.3. Statistics

Data for the sequences with /k/ were analyzed statistically using the General Linear Model Univariate analysis design (GENLIN) on IBM SPSS Statistics 20. Tests were run on mean values across tokens with ‘speaker’ as a random variable. The following ANOVAs were performed: two three-way ANOVAs on the F2 vowel transition frequency endpoints and ranges, with ‘vowel quality’ (‘i, a, u’), ‘consonant position’ (V#CV, VC#V, VCV, #CV and VC#, or a subset of these sequence structures depending on the test) and ‘vowel position’ (V1, V2) as factors; two three-way ANOVAs on the absolute and relative burst energy data with the factors ‘vowel quality’, ‘consonant position’ and ‘burst time’ (‘burst onset’, ‘burst midpoint’); three two-way ANOVAs on the closure duration data, and on the burst COG and duration values, with ‘vowel quality’ and ‘consonant position’ as independent factors. Bonferroni post-hoc tests were executed whenever independent variables with more than two levels turned out to be significant. Significant interactions were interpreted on the basis of results obtained from additional one-way ANOVAs run on all levels of a given factor while keeping the other factor constant. In all tests, the significance level was established at $p<0.05$.

Author's personal copy

2.2. Results

2.2.1. Vowel transitions

The initial hypothesis was that the vowel frequency component should contribute to velar softening as long as the F2 transition endpoint values match those for /tʃ/ above 1500 Hz, and the corresponding vowel transition ranges are large enough for the non-high vowels /a, u/. Moreover, an increase in F2 frequency could render [c] more affricate-like before /i/, and word and utterance initially (#CV, V#CV) than word and utterance finally (VC#, VC#V), the word medial intervocalic position being the least favored.

Statistical results for both the F2 transition endpoint values and ranges yielded a main effect of 'vowel quality' ($F(2, 46)=770.93$, $p<0.001$; $F(2, 46)=32.48$, $p<0.001$) and significant 'consonant position' ($F(4, 46)=22.47$, $p<0.001$; $F(4, 46)=16.11$, $p<0.001$), and 'consonant position' × 'vowel quality' interactions for each acoustic dimension ($F(8, 46)=24.05$, $p<0.001$; $F(8, 46)=12.21$, $p<0.01$). As shown in Fig. 1, these effects were associated with higher endpoint frequencies for /i/ (about 2000 Hz or higher) than for /a/ (about 1500–2000 Hz), and with higher endpoints for the (alveolo)palatal allophone [c] in the VC#V and VC# sequences (at 1350–1650 Hz) than for the velar allophone [k] in the V#CV, VCV and #CV sequences (below 1000 Hz) in the /u/ context condition. Thus, while the F2 frequency values for [c] appear to be appropriate for affricate identification in essentially all vowel contexts, they may be most prone to cue the affricate outcome in the context of /i/ and least prone to do so in the context of /u/. F2 endpoint differences as a function of position were marginally significant at the $p<0.05$ level for /i/ (V#CV > VCV) and /a/ (VCV > VC#), which is in support of a possible effect of articulatory reinforcement on velar softening in the word initial position in the /i/ vowel condition. The F2 vowel transition ranges are larger for /a/, and for /u/ in specific consonant positions (see below), than for /i/.

Some more relevant information about velar softening derives from the significant 'vowel position' × 'vowel quality' interactions ($F(2, 46)=8.85$, $p<0.001$; $F(2,46)=6$, $p<0.01$). Regarding the intervocalic sequences with /a/, F2 was significantly higher at V2 onset than at V1 offset and the CV transition ranges significantly greater than the VC ones, which is in accordance with the presence of a [j]-like configuration at closure offset and the possibility that it contributes to the generation of an affricate frication noise (see Fig. 1, middle graph and Introduction). As for those sequences where [c] is flanked by /u/, there was a higher F2 and larger VC transitions at V1 offset (1650 Hz for [uc#(u)], 1482 Hz for [uc#]) than at V2 onset (1371 Hz for [(u)c#u]), which suggests that there is more /u/-dependent lip rounding at the offset than at the onset of the (alveolo)palatal stop.

In sum, based on the vowel transitions data reported here and largely in agreement with Recasens and Espinosa (2009), the identification of [c] as /tʃ/ may be favored by the F2 transition endpoints in all contextual vowel conditions though presumably more so next to /i/ than to /a/ and next to /i, a/ than to /u/. Stop affrication should also be favored by an increase in the range of the VC and, more so, CV formant transitions in the /a/ context condition, and of the VC formant transitions in the /u/ context condition. Velar softening may also be triggered word initially when [c] occurs next to /i/.

2.2.2. Consonant period

As reported in Section 1.1, for the stop acoustic attributes to parallel those for /tʃ/, the burst should show at least a 3000 Hz spectral peak and be about 70 ms long or longer, while also approaching the intensity level of the adjacent vowels. In parallel to the vowel transitions frequency data, an increase in tongue contact associated with articulatory reinforcement could cause closure and burst duration, as well as burst spectral frequency, to increase word and utterance initially and less so word finally; on the other hand, an increase in burst intensity should render [c] most affricate-like intervocalically. Utterance final stops are expected to be especially long due to final lengthening.

2.2.2.1. Closure and burst duration. ANOVAs run on the stop closure duration data yielded a main effect of 'vowel quality' ($F(2, 22)=4.69$, $p<0.05$) which, as shown in Fig. 2, was associated with a slightly longer stop closure in the context of /u/ vs /a/ ($p<0.05$). Closure duration was also found to vary as a function of position ($F(3, 22)=41.83$, $p<0.001$), with the closing period being significantly longer utterance finally than elsewhere due to final lengthening (about 110 ms vs about 50-85 ms; $p<0.001$), but not word initially than word medially.

Regarding burst duration, the ANOVA yielded a main effect of 'vowel quality' ($F(2, 28)=4.80$, $p<0.05$) and 'consonant position' ($F(4, 28)=58.23$, $p<0.001$). As shown in Fig. 3, the burst was longer in the context of /i/ (50–70 ms) vs /a, u/ (less than 40 ms) across positional conditions, and differences between /i/ and /a/ reached statistical significance ($p<0.05$). This finding, which did not come out in Recasens and Espinosa (2009), suggests that /i/ may contribute to velar softening since the duration of the stop burst in this vowel context may parallel the minimal frication duration for /tʃ/. Moreover, analogously to the closure duration data, final lengthening caused the stop burst to be significantly longer utterance finally (115.5–122.1 ms) than in the remaining positions ($p<0.001$). Interestingly enough, data for the vowel /i/ in Fig. 3 also show some (non-significant) position-

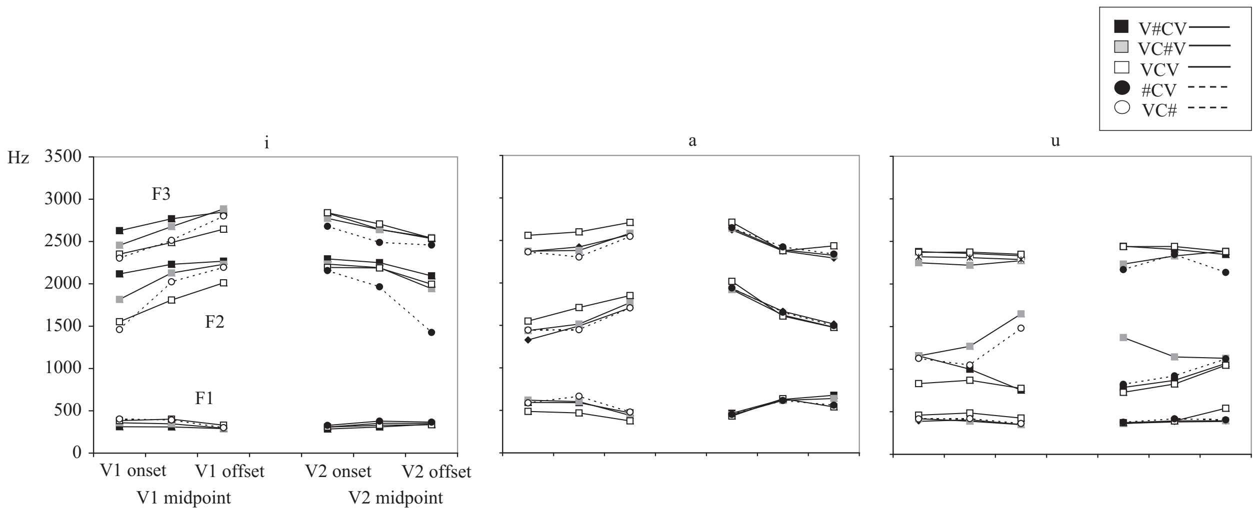

Fig. 1. Vowel formant trajectories for V#CV, VC#V, VCV, #CV and VC# sequences with /k/ and the vowels /i, a, u/. F1, F2 and F3 values are given at three temporal points during the vowel.

Author's personal copy

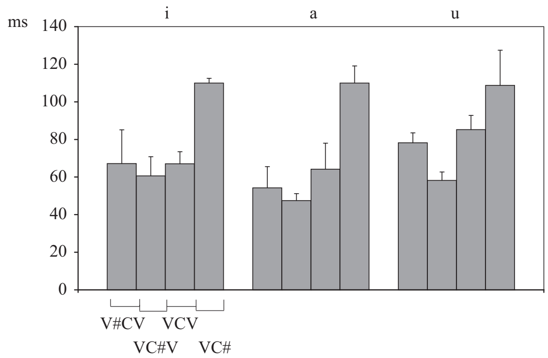

Fig. 2. Closure durations for /k/ as a function of contextual vowel (/i, a, u/) and sequence type (V#CV, VC#V, VCV, VC#). Error bars correspond to one standard deviation.

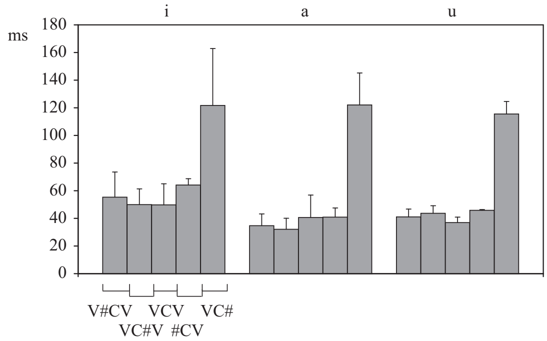

Fig. 3. Burst duration values for /k/ plotted as a function of the contextual vowels /i, a, u/ and the sequence types V#CV, VC#V, VCV, #CV and VC#. Error bars correspond to one standard deviation.

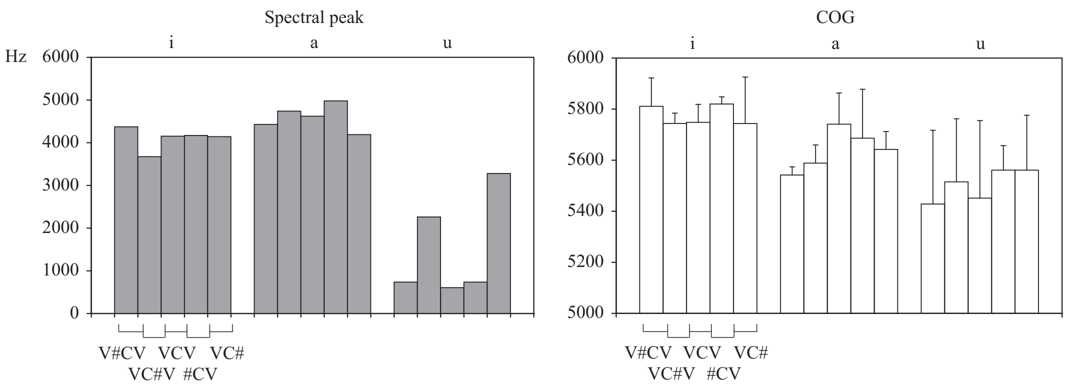

Fig. 4. Spectral peak and COG frequency values (left and right graph, respectively) for the /k/ burst plotted as a function of the contextual vowels /i, a, u/ and the sequence types V#CV, VC#V, VCV, #CV and VC#. Error bars correspond to one standard deviation.

dependent differences in stop burst duration varying in the progression /#ki/>/i#ki/>/ik#i, iki/, which suggests that the consonant may be reinforced by lengthening its stop burst word- and utterance-initially whenever situated next to the high front vowel.

2.2.2.2. Burst spectral characteristics. As shown by Fig. 4 (left graph), the burst spectral peak frequency for /k/ varies with vowel context and position. It is highest and thus affricate-like for the (alveolo)palatal allophone [c] in sequences with contextual /i, a/ (3500–5000 Hz) and somewhat higher next to /a/ (4189–4980 Hz) than next to /i/ (3675–4373 Hz), and there are no clear position-dependent effects on its frequency location in any of the two vowel contexts. Lip rounding and perhaps less dorsopalatal contact account for why [c] shows a lower spectral peak next to /u/, which turned out to be somewhat higher for /uk#/ (3280 Hz) than for /uk#u/ (2260 Hz) and therefore affricate-like in the former case but not in the latter. The lowest peak frequency corresponds to the velar allophone [k] in the sequences /u#ku, uku, #ku/ (606–735 Hz).

Another way to look for possible spectral differences among consonant realizations is through inspection of the overall burst spectral configurations. The left and middle graphs in Fig. 5 reveal similar cross-speaker spectral shapes for the [c] burst in the two vowel contexts /i/ and /a/. These spectral configurations exhibit a lower peak at the same frequencies indicated by the bars in the left graph of Fig. 4, a higher peak at about 10 kHz, and a flat plateau extending between the two peaks. Moreover, they show no apparent position-dependent differences in peak location and spectral shape though there appears to be a trend for [c] to exhibit a higher amplitude spectral configuration in utterance-initial position. The burst spectra for the (alveolo)palatal and velar allophones of /k/ in the context of the vowel /u/ plotted in the right graph of the figure exhibit peaks at similar

Author's personal copy

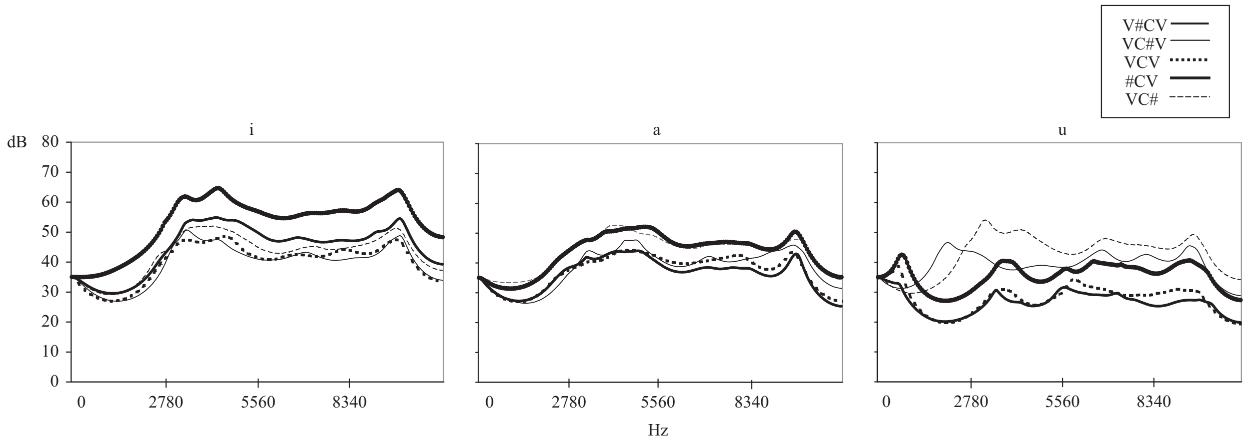

Fig. 5. LPC spectra for the /k/ burst plotted as a function of the contextual vowels /i, a, u/ and the sequence types V#CV, VC#V, VCV, #CV and VC#.

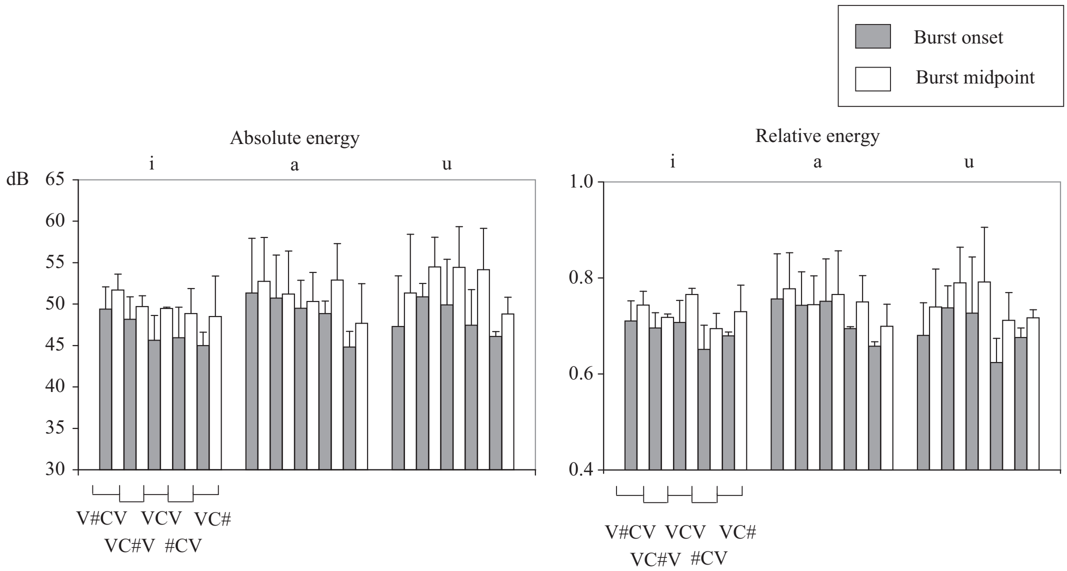

Fig. 6. Absolute and relative burst energy values for /k/ plotted as a function of the contextual vowels /i, a, u/ and the sequence types V#CV, VC#V, VCV, #CV and VC#. Filled and unfilled bars show energy values taken at burst onset and at burst midpoint, respectively. Error bars correspond to one standard deviation.

frequencies to those denoted by the bars in Fig. 4 (left graph), i.e., at 3280 Hz for [c] in the VC# sequence, at 2260 Hz for [c] in the VC#V sequence, and below 1000 Hz for [k] in the three remaining sequences V#CV, VCV and #CV.

Statistical tests run on the burst COG values yielded a main 'vowel quality' effect ($F(2, 28)=12.03$, $p<0.001$) which, as shown by the right graph of Fig. 4, was associated with a significantly higher value for /i/ (5743.4–5820.1) than for /u/ (5428.2–5693.4) and /a/ (5741.1–5541.9) at the $p<0.001$ and $p<0.05$ levels of significance, respectively. The COG data match the spectral peak frequency data in the left graph of the figure in showing no effect of consonant position and differs from it in exhibiting a higher value in the context of /i/ vs /a/.

In sum, the burst frequency component parallels the vowel transition endpoints in favoring [c] affrication in all vowels contexts and, assuming that affrication may increase with frequency height, next to /i/ followed by contextual /a/ and /u/ in this order. No clear preference for any specific position arises from the burst spectral data.

2.2.2.3. Burst energy. The left and right graphs in Fig. 6 plot absolute and relative burst energy values as a function of vowel context and consonant position at burst onset (filled bars) and at burst midpoint (unfilled bars).

Statistical results are similar for the two energy data types. As revealed by the left graph of the figure, the burst time effect was related to higher energy values at burst midpoint than at burst onset, i.e., unfilled bars are higher than filled bars, both for the absolute and relative energy data ($F(1, 58)=13.15$, $p<0.001$; $F(1, 58)=12.41$, $p<0.001$). On the other hand, the 'vowel' effect achieved significance for the absolute burst energy data only ($F(2, 58)=3.07$, $p<0.05$), and was associated with the presence of a slightly higher burst energy level for [c] before /u/ than before /i/ ($p<0.05$); absolute and relative burst energy values were also higher in the context of /a/ than of /i/ though this difference did not reach statistical significance. The 'consonant position' effect was significant for both energy types ($F(4, 58)=3.30$, $p<0.01$; $F(4, 58)=4.54$, $p<0.01$). As shown in the graphs, this significant effect was associated with a trend for the stop burst to be somewhat more intense intervocalically than utterance initially and utterance finally, i.e., for the V#CV, VC#V vs VC# sequences (absolute intensity), and for the V#CV, VCV vs #CV and the VCV vs VC# sequences (relative intensity). Based on these burst energy data, it may be hypothesized that [c] should be most prone to be identified as an affricate next to low and back rounded vowels, and in the intervocalic position perhaps mostly so when a word boundary is present.

Author's personal copy

2.2.3. Summary and hypothesis reevaluation

The analysis data presented in the preceding sections allow reevaluating the two hypotheses about the effect of vowel context and position on velar softening formulated in Section 1.2.

As for the effect of vowel context and in agreement with the most general contextual patterns of velar softening in the world's languages, the analysis results reported in the present investigation suggest that we ought to be in a better position to test Hypothesis 1 than in our previous study and, therefore, to elicit whether [c] may be categorized as an affricate more often before /i/ rather than before other vowels. This could be so since the vowel transition endpoint and burst spectral frequency values happen to be higher when [c] is followed by /i/ than when it precedes other vowels. Moreover, while in Recasens and Espinosa (2009) the stop burst was less than 40 ms long before any vowel, burst duration turns out to be higher before /i/ (50–70 ms) than before /a, u/ in the present investigation, thus meaning that the three Majorcan Catalan speakers SI, LLO and AN produced the sequence [ci] with a long burst frication noise which could be categorized as an affricate by listeners. Burst energy values and vowel transition ranges, on the other hand, were in agreement with Recasens and Espinosa (2009) in being higher and thus more affricate-like next to /a, u/ than to /i/.

As for the role of position, the vowel transition endpoint frequency and burst duration values suggest that velar softening in the /i/ context condition could be related to articulatory reinforcement word initially. As for the /a, u/ contexts, the analysis results indicate that [c] affrication may be cued by an increase in burst energy intervocalically rather than utterance initially and finally. Moreover, analogously to Recasens and Espinosa (2009), the stop closure and burst turned out to be especially long utterance-finally which may account for why prepausal [c] may be perceived as [ç] in this utterance position. Within this scenario, the prediction of Hypothesis 2 that the categorization of [c] as /tʃ/ should operate mostly in word-initial intervocalic position and, to a lesser extent, in word-final intervocalic position becomes feasible if we assume that the endpoint frequency of the vowel transitions and burst duration (which achieve their maximum word initially), and burst energy (which is highest intervocalically), contribute jointly to the affricate percept. The analysis data reported in the present investigation provide some support for this prediction in the case of contextual /i/ (see Sections 2.2.1 and 2.2.2.1), but less so in the context of /a/ where burst energy was highest intervocalically irrespective of word position.

3. Perception

In light of the analysis data presented in Section 2.2, this section reports results from two perception tests seeking to uncover the acoustic cues which cause [c] to be identified as /tʃ/ and to verify the hypotheses on velar softening formulated in Section 1.2. Test 1 explores the following issues using stimuli for a full set of vowel contexts and consonant positions: whether Catalan subjects are prone to identify [c] as /tʃ/ before /i/ rather than before /a, u/; the extent to which affricate identification percentages become especially high word-initial intervocalically and perhaps word-final intervocalically as well. Perception Test 2 has been designed in order to ascertain whether, in line with the author's auditory impression and the perceptual results obtained in Recasens and Espinosa (2009), [c] may be perceived highly often as /tʃ/ in the case of the sequence [uc#u]. A reason for this outcome, which does not match the conditions under which velar softening takes place in the world’s languages, should lie in the relatively high burst energy level and perhaps the large F2 frequency range of the VC transitions (about 400 Hz) rather than in the burst duration and spectral characteristics since [c] was found to exhibit a short burst and a not too high burst spectral peak frequency in the sequence [uc#u]. The issue as to why a candidate for stop affrication is not likely to be a candidate for velar softening will be addressed in the Discussion section.

3.1. Method

The perception stimuli of Tests 1 and 2 were composed of CV excerpts with /k/ excised from a subset of the V#CV, VC#V, VCV and #CV sequences subjected to analysis, which contained the stop burst and the vowel following the consonant. Sequences with utterance-final /k/ were not included in the perception tests since the (alveolo)palatal consonant sounded [ç]-like rather than [tʃ]-like in this case (see Section 1.1.2). Test 1 was composed of stimuli taken from speaker AN’s productions. It included five repetitions of three tokens of those sequences where /k/ exhibits an (alveolo)palatal stop realization, i.e., word-initial intervocalic /i#ki, a#ka/, word-final intervocalic /ik#i, ak#a, uk#u/, word-medial intervocalic /iki, aka/ and utterance-initial /#ki, #ka/. Five repetitions of three tokens of /#ku/, where /k/ is realized invariably as a velar stop, were also included in the test so as to achieve a better balance between the number of /k/ and /tʃ/ sounding stimuli. Test 2 was composed of excerpts of the sequences /#ku/ and /uk#u/ produced by speakers SI and LLO, i.e., five repetitions of three tokens of /uk#u/ where /k/ is realized as [c] and of three tokens of /#ku/ where /k/ is realized as a velar. The original data were split in two different tests in this way because informal perception of the CV excerpts showed that the [c] productions by speaker AN could be identified as /tʃ/ in several positions and vowel contexts, while those of the other speakers sounded /tʃ/-like only in the case of the sequence /uk#u/. A look at the production data of the three individual speakers in the /i/ context condition, and more specifically for /i#ki, ik#i, iki, #ki/ (a crucial case in the present investigation), reveals indeed the presence of more prominent affrication cues for speaker AN than for speakers SI and LLO. That was so for the F2 transition endpoints and, more importantly, for the burst duration ranges (61.4–72.6 ms, speaker AN; 48.2–59.2 ms, LLO; 35.2–68.4 ms, SI) and the burst intensity ranges (generally above 48.5 dB after the burst spike and 50 dB at the burst midpoint for speaker AR, and below these figures for the two other speakers). In fact, the burst for [c] before /i/ was longer for speaker AN than for any other of the nine speakers subject to recording (see Section 2.1.1). It appears that subject AN resembles more than subjects SI and LLO Majorcan Catalan speakers from older generations where [c] must have had a more strongly palatalized and affricate-like quality than it has today. Judging from the speaker-dependent analysis data reported in the present study and from the production data and perception results referred to in Recasens and Espinosa (2009), had we included acoustic stimuli for all three speakers in a single identification test the number of affricate responses would have probably been extremely low.

The total number of stimuli was 150 (3 tokens × 10 sequences × 5 repetitions × 1 speaker) in Test 1 and 60 (3 tokens × 2 sequences × 5 repetitions × 2 speakers) in Test 2. Stimuli were normalized for intensity. For this purpose, the energy level at the midpoint of the vowel following the consonant was measured for each original stimulus on energy contours using the same method described in Section 2.1. Then, the highest energy value of all stimuli was divided by the energy level of each stimulus, and the latter was multiplied by the resulting ratio. Moreover, the resulting energy values of all stimuli were increased by about 5 dB to ensure that stimuli with high vowels, mostly /i/, could be heard satisfactorily by the informants who took the perception tests.

In the two tests, there was a 3 s silence between consecutive stimuli and a 5 s silence every 10 stimuli. Tests 1 and 2 were played out to one or two informants per session using PowerPoint on a laptop with a good quality sound card through loudspeakers in quiet listening conditions. Fifteen subjects took Test 1 and 10 subjects took Test 2. A smaller number of informants were asked to take part in the second test since, as reported in

Author's personal copy

Section 3.2 and in line with the author's auditory impressions, all 10 subjects were highly consistent in categorizing the [k]-sounding stimuli as /k/ and the [c] sounding stimuli as /tʃ/.

Subjects were asked to identify the stimuli as /k/ or /tʃ/ and to write down 'K' or 'TX' (which is the digraph for the affricate in Catalan) on an answer sheet. They all speak the Eastern Catalan dialect. Though the Eastern dialect of Catalan does not have the [c] allophone of /k/ in its sound inventory (see Section 1), speakers of this dialect are familiar with it because they hear it often from Majorcan Catalan speakers studying and/or living in Catalonia. This applies all the more so in our case since the informants who took the two tests were almost exclusively university professors and university students. Moreover, the fact that the Majorcan and Eastern Catalan subjects who took the perception test in Recasens and Espinosa (2009) yielded highly similar affricate identification percentages as a function of vowel context and position suggests that speakers of the two dialectal origins perceive the stop in analogous ways.

The perceptual identification data for Test 1 were submitted to a logistic regression analysis using the Generalized Linear Mixed Model procedure (GENLINMIXED) with IBM SPSS Statistics 20. The variable 'sequence' was entered as a fixed factor (levels 'i#ki, a#ka, ik#i, ak#a, uk#u, iki, aka, #ki, #ka') and 'subject' as a random factor. Data for /#ku/ were excluded from the statistical analysis since this sequence yielded no /tʃ/ responses. 'Vowel' ('i, a, u') and 'consonant position' ('V#CV, VC#V, VCV, #CV') were not entered as independent variables since data for the contextual /u/ condition were processed for /uk#u/ only. Pairwise comparisons between levels of the 'sequence' factor were performed using the Bonferroni adjustment. The significance level was established at $p<0.05$.

In order to ascertain which acoustic characteristics acted as best affricate identification cues in the subjects' responses, bivariate correlation analyses were performed between the /tʃ/ identification percentages obtained in Test 1 and data for the following ten analysis parameters: VC and CV transitions endpoint frequency; VC and CV transitions ranges; closure and burst duration; burst spectral peak and COG; absolute and relative energy at the burst midpoint. Mean values across tokens for those productions of the sequences /i#ki, a#ka, ik#i, ak#a, iki, aka, #ki, #ka/ which were submitted to perceptual evaluation were entered in the correlation analyses. Data for /uk#u/ were not included since, as shown in Section 3.2, this sequence behaved differently from all sequences with the vowels /i, a/ in important respects. In addition, a multivariate linear regression analysis was carried out on the same perception data in order to find out which analysis variables could act as predictors of the affricate identification responses. The significance level was established at $p<0.05$ for both correlation and regression tests.

3.2. *Results*

3.2.1. *Test 1*

The affricate identification percentages for Test 1 yielded a significant 'sequence' effect ($F(8, 396)=35.02$, $p<0.001$). As revealed by the pairwise comparisons and as shown in Table 1 and Fig. 7 (left graph), the identification of [c] as /tʃ/ occurred most often in the case of the sequences /uk#u, i#ki, ik#i/ in this order. As for the remaining sequences, /aka/ was least favored (results for this sequence differed significantly from those for /iki, #ki, ak#a/), and there was also a significant effect for /iki/>/#ka/.

An interpretation of these identification results with reference to the two hypotheses formulated in Section 1.2 follows. There is a clear preference for the affricate to be heard in sequences with /i/ rather than in those with /a/, as revealed by significant differences in affricate identification for the following pairs of sequences sharing the same positional condition: /i#ki/>/a#ka/; /ik#i/>/ak#a/; /iki/>/aka/. Regarding position, there is also a clear trend for /tʃ/ to be identified in word-initial and word-final intervocalic position rather than utterance initially or word-medial intervocalically in sequences with /i/, i.e., the position hierarchy resulting from the statistical results is /i#ki, ik#i/>/#ki, iki/; moreover, even though non-significantly, affricate percentages for /i#ki/ exceeded those for /ik#i/. As for the /a/ condition, affricate identification percentages were higher when /k/ occurred next to a word or utterance boundary than when it showed up in the word-medial intervocalic position, i.e., affricate percentages were roughly the same for /a#ka, ak#a, #ka/ and significantly higher for /ak#a/ than for /aka/. The sequence /uk#u/ yielded a very high number of affricate responses.

Fig. 7 (right graph) shows that speakers differed quite dramatically in the way they perceived the acoustic stimuli, with the affricate identification percentages ranging from about 10% to 90% depending on the subject. In spite of these large affricate identification differences, informants were very consistent in exhibiting the sequence-dependent pattern of responses presented in the left graph of the figure and, therefore, in hearing /tʃ/ most often in the case of the sequence /uk#u/ and the sequences with /i/ in the progression mentioned above followed by most sequences with the vowel /a/. This scenario is indicated by the relatively small standard deviation values appended to most bars in the graph.

3.2.2. *Test 2*

As shown by Fig. 8 (left graph), results for the perception Test 2 were as expected, i.e., the percentage of affricate responses was null for /#ku/ and about 95% for /uk#u/ independently of whether the original stimuli were taken from productions by speaker SI or by speaker LLO. Moreover, data in Fig. 8 (right graph) reveal that subjects were highly consistent in perceiving /k/ when the stop occurred word utterance initially (about 50% of the stimuli) and /tʃ/ when it occurred in word-final intervocalic position (remaining 50%).

Table 1
Significant differences in /tʃ/ identification between pairs of sequence types ($^{*}p<0.05$, $^{**}p<0.01$, $^{***}p<0.001$).

	i#Ci	a#Ca	iC#i	aC#a	uC#u	iCi	aCa	#Ci	#Ca
i#Ci		***		***		***	***	***	***
a#Ca			***		***				
iC#i				***		**	***	***	***
aC#a					***		*		
uC#u						***	***	***	***
iCi							**		*
aCa								*	
#Ci									
#Ca									

Author's personal copy

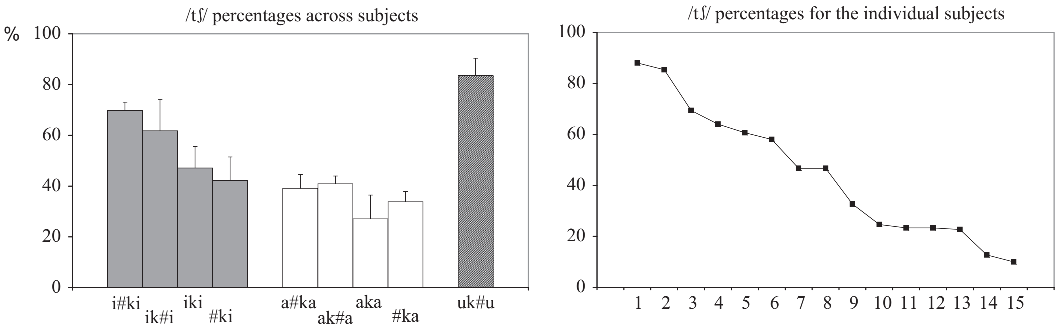

Fig. 7. Results for the perception Test 1. Percentages of /tʃ/ identification as a function of different sequence types (left) and of informants 1 through 15 who took the perception test (right). Error bars correspond to one standard deviation.

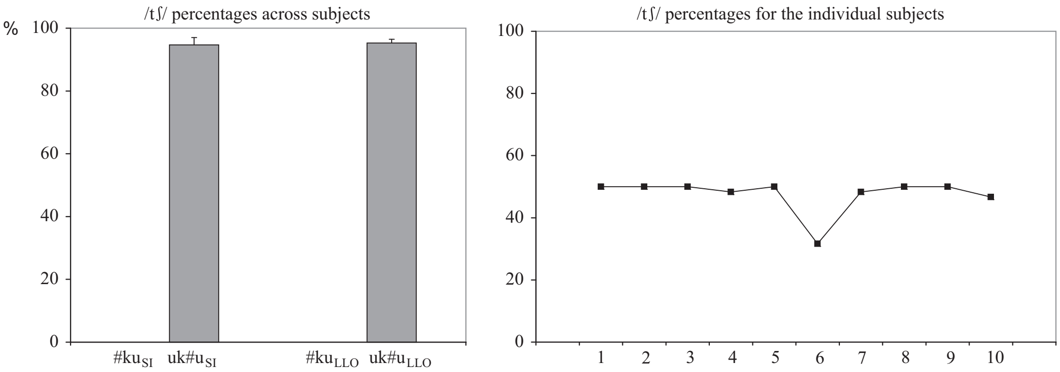

Fig. 8. Results for the perception Test 2. Percentages of /tʃ/ identification as a function of different sequence types (left) and of informants 1 through 10 who took the perception test (right). Error bars correspond to one standard deviation. Identification percentages across subjects are plotted separately for the sequence productions of speakers SI and LLO.

Table 2
Pearson correlation coefficients and significance level for correlation analyses between percentages of /tʃ/ identification and analysis parameter values (*p<0.05, **p<0.01, ***p<0.001).

	Pearson's correlation value	Significance level
VC frequency endpoint (F2)	0.855	***
CV frequency endpoint (F2)	0.866	***
VC frequency ranges	−0.559	
CV frequency ranges	−0.771	**
Closure duration	−0.405	
Burst duration	0.649	
Burst spectral peak	−0.006	
Burst COG	0.67	
Burst absolute energy	0.565	
Burst relative energy	0.171	

3.2.3. Correlations and regression between the analysis and perception values

Table 2 presents the results for the correlation analyses between the /tʃ/ identification percentages and the analysis parameters. According to the table, *r* values exceed 0.6 and are thus high and positive for four out of the 10 correlation pairs, i.e., those involving the endpoint frequency of the VC and CV transitions (which achieve significance), burst duration and burst COG.

Fig. 9 shows that there is indeed a good positive relationship between the affricate identification percentages (top graph), and the endpoint frequency of the VC and CV transitions and the burst duration and COG values (three middle graphs), in sequences where [c] occurs next to /i/ and /a/. Analogously to the analysis values for the speakers SI, LLO and AN reported in Figs. 1, 3 and 4, those plotted in Fig. 9 are higher for /i/ than for /a/ thus suggesting that all three analysis parameters may contribute to affricate identification to greater or lesser extent. These high and positive correlations appear to be also related to consonant position in the case of the /i/ context condition: as shown in the figure, the vowel transition endpoint frequency and burst duration values are highest for /i#ki/ and, less so, /ik#i/ (and thus word initial and word final intervocalically) and lowest for /iki/ and /#ki/ (and thus word medial intervocalically and in absolute word initial position). The frequency component and to a large extent burst duration are however not correlated with the affricate identification percentages for the sequence /uk#u/: while identification percentages for this sequence are about 80%, its frequency values are low and the stop burst is relatively short. In sum, the frequency characteristics of the vowel transitions and the stop burst, as well as burst duration, appear to account for differences in affricate identification between /i/ and /a/ and among positions in the case of the /i/ vowel condition.

The role of the vowel transition ranges and burst energy in affricate identification differs in important respects from that of the frequency and duration parameters referred to above. According to Table 2 and the left bottom graph of Fig. 9, correlation values between the affricate identification

Author's personal copy

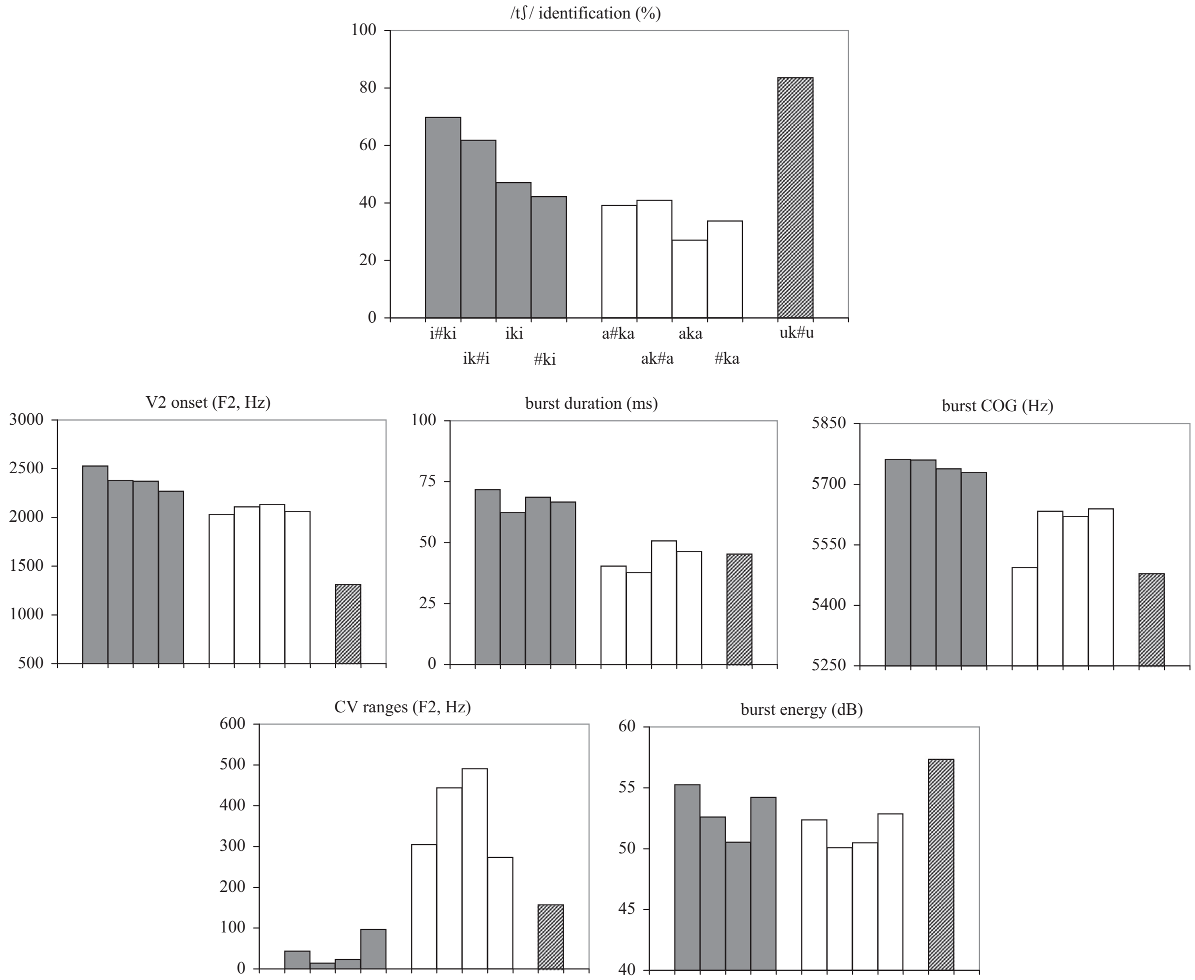

Fig. 9. (Top panel) Affricate identification percentages obtained in Test 1. (Middle and bottom panels) Analysis parameter values for the stop excerpts that were subjected to perceptual evaluation.

percentages and the CV transition ranges (also the VC transition ranges to a lesser extent) are high and negative since the ranges in question are much larger for /a/ than for /i/. It may thus very well be that this specific acoustic dimension contributes to the identification of /k/ as an affricate in the /a/ context. On the other hand, correlation values between affricate identification and burst absolute energy are positive and lower than those obtained for the CV endpoint frequencies and burst duration and COG. According to Fig. 9 (right bottom graph), the analysis values for this parameter appear to contribute mostly to the high affricate identification percentages for the sequence /uk#u/. This is so since the perceptual identification percentages and the burst energy values for this sequence exceed those for all other sequences subjected to perceptual evaluation. It should also be noticed that, in contrast with the burst energy data across speakers plotted in Fig. 6, the stop burst energy values for the perception stimuli presented in Fig. 9 are not higher for the /a/ context than for the /i/ context; in agreement with the affricate identification results and with Hypothesis 2 though, the burst energy values for both contextual conditions /a/ and /i/ are lower in the word medial than in the word initial intervocalic position.

The mutivariate regression model including only the significant predictors, i.e., the frequency endpoints and ranges of the vowel transitions (see also Table 2), yielded an adjusted *R*-squared of 0.67. While these two predictors turned out to contribute significantly to the affricate responses according to a one-way ANOVA ($F(2, 5)=8.09$, $p<0.05$), their regression coefficients could not be estimated reliably due to multicollinearity problems (there was a variance inflation factor of 2.87).

3.3. Summary and discussion

A major difference between the perception results reported in the present study and those reported in Recasens and Espinosa (2009) concerns the perceptual role played by the contextual vowels in the identification of [c] as /tʃ/: consistent with the predictions of Hypothesis 1 and in accordance with the universal pattern of velar softening, the presence of a relatively long burst and of especially high vowel transition endpoint and burst frequency characteristics in sequences with contextual /i/ rendered the affricate identification percentages highest in this vowel context. In addition to the spectral characteristics of the vowel transitions and the stop burst, the vowel transition ranges (and perhaps burst energy whenever data for all speakers subjected to analysis are taken into account) probably play a significant role in affricate identification for [c] before /a/. The high affricate identification percentages for the sequence [uc#u] in the two perception tests appear to be associated mainly with burst energy and perhaps with the VC transitions ranges as well.

Results for the effect of consonant position agree in many respects with the predictions formulated by Hypothesis 2. Perception results for contextual /i/ parallel those obtained in our previous paper where affricate identification percentages turned out to be higher intervocalically than utterance initially presumably since the presence of two adjacent vowels enhance the burst frication noise by raising its energy level. A novel outcome of the present investigation is that, consistent with the role of articulatory reinforcement in strong positions, affricate percentages for intervocalic [c] turned out to be higher word initially and word finally than word medially, and to some extent word initially than word finally.

Author's personal copy

Therefore, it appears that both stop placement next to a word boundary and the presence of two flanking vowels jointly enhance the affricate quality of the (alveolo)palatal stop. The perceptual and analysis data for the /a/ context appear to be also consistent with this effect of position on velar softening in several respects.

4. General discussion

The analysis data and perception results reported in the present investigation conform better than Recasens and Espinosa (2009) to the contextual patterns of velar softening observed in the world's languages while throwing new light on the phonetic causes of this sound change.

In contrast with Recasens and Espinosa (2009) and in agreement with the universal patterns of velar softening (and with Hypothesis 1), affricate percepts occurred mostly before /i/, where the spectral frequency characteristics are specially high and the burst frication noise may lengthen considerably as lingual constriction narrowing increases. To a lesser extent and also in agreement with the patterns of sound change, the stop may also be heard as an affricate before /a/ presumably since the vowel transition ranges and burst intensity may become especially prominent in this vowel context. This contextual hierarchy implies that, if velar softening applies before /a/, it should operate before front vocalic segments as well, while the reverse does not hold necessarily. Maximal affricate identification percentages for word final [c] in the /u/ context appear to be a special case and thus require a specific justification (see below), among other things since /k/ is realized as [k] before this vowel in all the other word and sentence positions in Majorcan Catalan.

Regarding position, the analysis and perceptual data lead support to the hypothesis that the word-initial intervocalic position is the best environment for the identification of [c] as an affricate, followed by the word-final intervocalic position (Hypothesis 2). It has been argued that this result is associated with two factors acting synergistically: on the one hand, closure and burst lengthening and some additional frequency rise at the vowel transitions endpoint, which appear to be related to articulatory reinforcement at the edges of words; on the other hand, an increase in burst intensity, which occurs when the stop is preceded and followed by a vowel. The perception section of our study provides support for the former aspect rather than for the latter since the burst energy data for the stimuli subjected to perceptual identification differed in several respects from those obtained across speakers. To the extent that they do not exhibit all the crucial acoustic attributes, the utterance initial and word-medial intervocalic positions appear to be less likely stop affrication candidates.

Additional perceptual testing is needed in order to ascertain the hierarchy of acoustic characteristics which cue the palatoalveolar affricate in specific contextual and positional conditions: whether the spectral component, burst duration or both are the major acoustic velar softening cues in the /i/ context; whether velar softening in the /a/ context is cued mainly by the vowel transition ranges or by the burst energy level; which one of these acoustic cues contributes most to velar softening in specific positions rather in others. Manipulation of individual acoustic parameters using synthetic speech appears to be the most appropriate methodological tool for exploring this issue. A preliminary suggestion based on the findings reported in this paper is that, given that the frequency characteristics of the vowel transitions and the stop burst were affricate-like in most vowel contexts, [c] affricate identification could depend on burst duration and energy to a large extent. This possibility is in accordance with the notion that affricates may be considered as stops with a long and intense frication period, and with the prominent role of noise duration and amplitude rise time in the discrimination between fricatives and affricates (Dorman, Raphael, & Isenberg, 1980).

In light of the position-dependent results reported in the present study, velar softening ought to be most prone to apply in the V#CV and VC#V environments, and the affricate outcome could extend from here to the postpausal and prepausal positions (see Section 1.1.2). It thus appears that the replacement of [c] by an affricate in word-initial position does not start out utterance initially, apparently since affrication depends not only on articulatory strengthening through closure and burst lengthening but also on an increase in burst intensity. This mechanism is in contrast with other changes which occur in strong positions and depend solely on a gain in tongue contact degree and in segmental duration such as the palatalization of dentals and alveolars (/n/>/ɲ/, as in [ˈɲaβo] for [ˈnaβo] “turnip” in Asturian Spanish) or the occlusivization of approximants (/j/>[ɟ], as in dialectal Spanish [ˈ[ɟ]uɣo] for [ˈjuɣo] “yoke”). Fricative outcomes may arise from [c] utterance finally where the burst is especially long, as well as in the weaker word-medial intervocalic position where the stop may become less salient. It remains unclear whether the shift of [c] into [ʃ] in the latter environment occurs through direct lenition (analogously to stop fricativization, as for /k/>[x, h], /p/> [ɸ] and /t/>[θ] in the Gorgia Toscana; Villafaña Dalcher, 2006), or else through closure reduction of an intermediate affricate outcome.

The position-dependent data reported in the present study are also relevant regarding the issue as to whether stop affrication should be considered a case of lenition and thus the first step for a voiceless stop to become a fricative (Lass, 1984: 108), or else an instance of fortition through which stops acquire a large tongue contact area and a longer and more intense frication period when changing into affricates (Buiza & Plug, 2012, Lavoie, 2001). Our data show that the identification of [c] as [tʃ] involves segmental reinforcement in strong positions such as word initially vis-à-vis weaker ones such as word-medial intervocalically (Ségéral & Scheer, 2008, Szigetvári, 2008). This view is also consistent with Garrett and Johnson (2010) hypothesis that velar softening may be achieved through phonetic enhancement. These authors distinguish at least two enhancement types depending on whether a new articulatory feature is introduced or not: articulatory enhancement, as for the change [uCi]>[yCi] resulting from a context-dependent anterior realization of the high back vowel; perceptual enhancement, as for the labalized realization [x^w] of [x] causing the fricative consonant to be perceived as /f/ (e.g., Middle English *lauxe*>*laugh*). The data reported in the present investigation suggest that velar softening may be achieved through articulatory enhancement since the acoustic seed for the change is available at the (alveolo)palatal stop release and should be associated with specific articulatory maneuvers and aerodynamic factors such as an increase in lingual constriction narrowing and of airflow volume passing through the constriction.

A comment needs to be added about the high affricate identification percentages for the sequence [uc#u]. In principle, this case should not be considered an exception since the high burst energy level and the large frequency range of the vowel transitions ought to render [c] affrication feasible in the sequence in question. What appears to be striking is that a realization yielding 80% of affricate percepts is not a good candidate for velar softening in the world's languages. In order to account for this anomaly, we would like to propose that the high affricate percentages for [uc#u] may have been associated with the fact that there was not enough acoustic information in the short perception stimuli for listeners to identify the surrounding vowel as /u/. It may very well be that, when listeners are confronted with the full sequence [uc#u], factors such the grave quality of the two vowels and implicit knowledge about /k/ being realized as [k] before /u/ word initially and word medially prevent [c] from being identified as /tʃ/. In these circumstances, listeners may fail to categorize the affricate-like information contained in the (alveolo)palatal stop burst as part of a new affricate phoneme.

Author's personal copy

Data reported in this paper are also in support of the notion that specific speakers and listeners may initiate a given sound changes rather than others (Grosvald & Corina, 2012; Harrington, Kleber, & Reubold, 2012). As for the speakers' role, it should be mentioned that subjects may differ a great deal both with respect to dorsopalatal contact size and closure location for (alveolo)palatal stops (Recasens & Espinosa, 2006), as well as to stop burst frication degree before high front vowels or glides. As for the role of the listener, the affricate identification percentages differed considerably among the informants who took the perception test. Other studies also report that subjects may be more or less sensitive to the role that specific acoustic cues play in the substitution of one phonetic sound by another, as for example regarding the perceptual effectiveness of the vowel formant transitions and the steady-state formants during the consonant period in the replacement of dark /l/ by /w/ (Recasens & Espinosa, 2010a).

Analysis and perception data reported in this investigation reveal that, at least for unaspirated stops, velar softening may be triggered by considerably anterior (alveolo)palatal realizations of the velar stop. Since these articulatory realizations (also alternations between [c] and [tʃ]) are widely available in the Romance languages where velar softening has occurred, it seems plausible to hypothesize that at least in this linguistic domain the sound change process in question has been triggered by (alveolo)palatal stop productions of /k/. The same rationale could be applied to languages such as Czech or Hungarian where unaspirated [c] has phonemic status and in Greek where the stop is allophonic as in Majorcan Catalan (see Introduction section).

This study has also demonstrated that the production and perception mechanisms which give rise to a palatoalveolar affricate out of an (alveolo) palatal stop may be replicated in the laboratory. Several factors appear to favor the implementation of velar softening: the presence of a relatively prominent (alveolo)palatal stop burst and high frequency F2 vowel transition endpoints whenever the consonant is followed mostly by a high front vowel and occurs in word-initial and word-final intervocalic position. Moreover, speakers may differ quite considerably as to whether they are able to produce affricate-like realizations of [c] and to identify the (alveolo)palatal stop as /tʃ/ in the relevant contextual and positional conditions. This study also suggests that favorable cues to velar softening may not be used necessarily by listeners, as for those available in the sequence /uk#u/ ([uc#u]) where the implicit knowledge of velar realizations in other word positions could refrain the affricate percept from being categorized phonemically.

Acknowledgments

This research has been funded by the project FFI-2009-09339 from the Ministry of Innovation and Science of Spain, by the Catalan Institution for Research and Advanced Studies (ICREA), and by the research group 2009SGR003 from the Generalitat de Catalunya. I would like to thank four anonymous reviewers for comments, and Juli Cebrian, Jaume Corbera and Pilar Prieto for their help with the informant selection process.

Appendix A

List of sentences in Catalan orthography and with English glosses. The segmental sequences subjected to analysis are given in phonetic transcription and appear underlined in the Catalan orthographic representation. In sentence 2, the word-final rhotic has been enclosed within parentheses because it is not pronounced in Majorcan Catalan.

V#CV	1. [ici]	repartí quissos	"he delivered dogs"
	2. [aca]	va porta(r) caspa	"he/she had dandruff"
	3. [uku]	ell li duu cusses	"he brings female dogs for him"
VC#V	4. [ici]	un bonic indi	"a handsome Indian"
	5. [aca]	és un sac aspre	"this is a rough sack"
	6. [ucu]	és un suc útil	"this is a useful juice"
VCV	7. [ici]	ell s'ha enriquit	"he has enriched himself"
	8. [əca]	ell s'ha tacat	"he got stained"
	9. [oku]	és un bocut	"he has a big mouth"
#CV	10. [ci]	quilos d'or	"golden kilos"
	11. [ca]	casa baixa	"a low house"
	12. [ku]	cus la roba	"he/she sows the cloth"
VC#	13. [ic]	l'avi és ric	"the grandfather is rich"
	14. [ac]	umple es sac	"fill the sack"
	15. [uc]	treu el suc	"take out the juice"

References

Barnes, J. (2006). *Strength and weakness at the interface: Positional neutralization in phonetics and phonology*. Berlin/NewYork: Mouton de Gruyter.
Bhat, D. (1978). A general study of palatalization. In: J. Greenberg (Ed.), *Universals of human language, volume 2, Phonology* (pp. 47–92). Stanford: Stanford University Press.
Blumstein, S. E. (1986). On acoustic invariance in speech. In: J. Perkell, & D. Klatt (Eds.), *Invariance and variability in speech processes* (pp. 178–193). Hillsdale: Erlbaum.
Buizza, E., & Plug, L. (2012). Lenition, fortition and the status of plosive affrication of spontaneous RP English /t/. *Phonology*, *29*, 1–38.
Chang, S. S., Plauché, M. C., & Ohala, J. J. (2001). Markedness and consonant confusion asymmetries. In: E. Hume, & K. Johnson (Eds.), *The role of speech perception and phonology* (pp. 79–102). San Diego and London: Academic Press.
Cho, T., Jun, S.-A., & Ladefoged, P. (2002). Acoustic and aerodynamic correlates of Korean stops and fricatives. *Journal of Phonetics*, *30*, 193–228.
Corneau, C., Soquet, A., & Demolin, D. (2000). Une étude électropalatographique de la palatalisation des occlusives vélaires en français. *Actes des XXIIIèmes Journées d'Études sur la Parole*, Aussois, France, 445–448.
Delattre, P. C., Liberman, A. M., & Cooper, F. S. (1955). Acoustic loci and transitional cues for consonants. *Journal of the Acoustical Society of America*, *27*, 769–773.

Author's personal copy

Docherty, G. (1992). *The timing of voicing in British English obstruents*. Berlin/New York: Foris Publications.
Dorman, M. F., Raphael, L., & Isenberg, D. (1980). Acoustic cues for a fricative-affricate contrast in word-final position. *Journal of Phonetics*, *8*, 397–405.
Dorman, M. F., Studdert-Kennedy, M., & Raphael, L. J. (1977). Stop-consonant recognition: Release bursts and formant transitions as functionally equivalent, context-dependent cues. *Perception and Psychophysics*, *22*, 109–122.
Elwert, W. T. (1943). *Die Mundart des Fassa-Tals*. Heidelberg: C. Winter.
Fant, G. (1960). *Acoustic theory of speech production*. The Hague: Mouton.
Fougeron, C., & Keating, P. A. (1997). Articulatory strengthening at edges of prosodic domains. *Journal of the Acoustical Society of America*, *101*, 3728–3740.
Garrett, A., & Johnson, K. (2010). Phonetic bias in sound change. In: A. C. Yu (Ed.), *Origins of Sound Change: Approaches to Phonologization* (pp. 51–79). Oxford: Oxford University Press.
Grosvald, M., & Corina, D. P. (2012). The production and perception of sub-phonemic vowel contrasts and the role of the listener in sound change. In: M. J. Solé, & D. Recasens (Eds.), *The initiation of sound change. Perception, production, and social factors* (pp. 77–102). Amsterdam: John Benjamins.
Guion, S. G. (1996). *Velar palatalization, coarticulation, perception and sound change (Ph.D. dissertation)*. Austin: University of Texas.
Guion, S. G. (1998). The role of perception in the sound change of velar palatalization. *Phonetica*, *55*, 18–52.
Harrington, J., Kleber, F., & Reubold, U. (2012). The production and perception of coarticulation in two types of sound changes in progress. In: S. Fuchs, M. Weirich, D. Pape, & P. Perrier (Eds.), *Speech planning and dynamics* (pp. 33–55). Frankfurt: Peter Lang.
Keating, P., & Lahiri, A. (1993). Fronted velars, palatalized velars, and palatals. *Phonetica*, *50*, 73–101.
Keating, P., Wright, R., & Zhang, J. (1999). Word-level asymmetries in consonant articulation. *UCLA Working Papers in Phonetics*, Vol. 97 (pp. 152–157).
Ladefoged, P. (2003). *Phonetic data analysis. An introduction to fieldwork and instrumental techniques*. Malden, MA: Blackwell.
Lass, R. (1984). *Phonology: An introduction to basic concepts*. Cambridge: Cambridge University Press.
Lausberg, H. (1970). *Lingüística románica*. Madrid: Gredos.
Lavoie, L. A. (2001). *Consonant strength: Phonological patterns and phonetic manifestations*. New York: Garland Publishing, Inc.
Liberman, A., Delattre, P. C., Cooper, F. S., & Gerstman, L. J. (1954). The role of consonant–vowel transitions in the perception of the stop and nasal consonants. *Psychological Monographs*, *68*, 1–13.
Luzi, J. (1904). Der sutselvischen Dialekte. *Romanische Forschungen*, *16*, 757–846.
Maddieson, I. (1984). *Patterns of sounds*. Cambridge: Cambridge University Press.
McDonough, J., & Wood, V. (2008). The stop contrast of the Athabaskan languages. *Journal of Phonetics*, *36*, 427–449.
Mielke, J. (2007). P-base. 〈http://aix1.uottawa.ca/∼jmielke/pbase〉.
Ohala, J. J. (1992). What's cognitive, what's not, in sound change. In: G. Kellermann, & D. Morrisey (Eds.), *Diachrony within synchrony: Language history and cognition* (pp. 309–335). Frankfurt: Peter Lang Verlag.
Pieras, F. (1999). *Social dynamics of language contact in Palma de Mallorca* (Ph.D. thesis). The Pennsylvania State University.
Pope, M. K. (1934). *From Latin to modern French*. Manchester: Manchester University Press.
Recasens, D. (1986). *Estudis de fonètica experimental del català oriental central*. Barcelona: Publicacions de l'Abadia de Montserrat.
Recasens, D. (1991). Timing in Catalan. *Actes du XIIème Congrès International des Sciences Phonétiques*, Vol. 4. Université de Provence, 230–233.
Recasens, D. (2013). On the articulatory classification of (alveolo)palatal consonants. *Journal of the International Phonetic Association*, *43*, 1–23.
Recasens, D., & Espinosa, A. (2006). Articulatory, positional and contextual characteristics of palatal consonants: Evidence from Majorcan Catalan. *Journal of Phonetics*, *34*, 295–318.
Recasens, D., & Espinosa, A. (2009). Acoustics and perception of velar softening for unaspirated stops. *Journal of Phonetics*, *37*, 189–211.
Recasens, D., & Espinosa, A. (2010a). The role of spectral and temporal cues in consonantal vocalization and glide insertion. *Phonetica*, *67*, 1–24.
Recasens, D., & Espinosa, A. (2010b). Lingual kinematics and coarticulation for alveolopalatal and velar consonants in Catalan. *Journal of the Acoustical Society of America*, *127*, 3154–3165.
Recasens, D., & Pallarès, M. D. (2001). *De la fonètica a la fonologia*. Barcelona: Ariel.
Remacle, L. (1953). *Atlas linguistique de la Wallonie, vol. 1 (Introduction générale, aspects phonétiques)*. Liège: Université de Liège.
Rohlfs, G. (1966). *Grammatica storica della lingua italiana, vol. 1, Fonetica*. Turin: Einaudi.
Ségéral, P., & Scheer, T. (2008). Positional factors in lenition and fortition. In: J. Brandao de Carvalho, T. Scheer, & P. Ségéral (Eds.), *Lenition and fortition* (pp. 131–171). Berlin: Mouton de Gruyter.
Šimáčková, Š., Podlipský, V. J., & Chládková, K. (2012). Czech spoken in Bohemia and Moravia. *Journal of the International Phonetic Association*, *42*, 225–232.
Simonet, M. (2010). Dark and clear laterals in Catalan and Spanish: Interaction of phonetic categories in early bilinguals. *Journal of Phonetics*, *38*, 664–679.
Stevens, K. N. (1998). *Acoustic phonetics*. Cambridge, USA: The MIT Press.
Syrika, A., Kong, E. J., & Edwards, J. (2011). Velar softening: An acoustic study in Modern Greek. In: W.-S. Lee, & E. Zee (Eds.), *Proceedings of XVII ICPhS* (pp. 1926–1929). Hong Kong: City University of Hong Kong.
Szigetvári, P. (2008). What and where? In: J. Brandao de Carvalho, T. Scheer, & P. Ségéral (Eds.), *Lenition and fortition* (pp. 93–129). Berlin: Mouton de Gruyter.
Veny, J. (1978). *Els parlars catalans*. Palma de Mallorca: Raixa.
Villafaña Dalcher, C. (2006). *Consonant weakening in Florentine Italian. An acoustic study of gradient and variable sound change* (Ph.D. thesis), Georgetown University, Washington.
Zygis, M., Recasens, D., & Espinosa, A. (2010). The role of position in testing the acoustic-equivalence hypothesis of velar softening for aspirated stops. In: S. Fuchs, P. Hoole, C. Mooshammer, & M. Zygis (Eds.), *Between the regular and the particular in speech and language* (pp. 125–144). Frankfurt: Peter Lang.