languages

Article

The Effect of Manner of Articulation and Syllable Affiliation on Tongue Configuration for Catalan Stop–Liquid and Liquid–Stop Sequences: An Ultrasound Study

Daniel Recasens

Departament de Filologia Catalana, Universitat Autònoma de Barcelona, 08193 Bellaterra, Barcelona, Spain; daniel.recasens@uab.es

Abstract: The present study reports tongue configuration data recorded with ultrasound for two sets of consonant sequences uttered by five native Catalan speakers. Articulatory data for the onset cluster pairs [kl]-[ɣl] and [kɾ]-[ɣɾ], and also for [l#k]-[l#ɣ] and [r#k]-[r#ɣ], analyzed in the first part of the investigation revealed that, as a general rule, the (shorter) velar approximant is less constricted than the (longer) voiceless velar stop at the velar and palatal zones while exhibiting a more retracted tongue body at the pharynx. These manner of articulation-dependent differences may extend into the preceding liquid. Data for [k#l]-[kl] and [k#r]-[kɾ] dealt with in the second part of the study show that the velar is articulated with more tongue body retraction for [k#l] vs. [kl] and for [k#r] vs. [kɾ], and with a higher tongue dorsum for [k#l] vs. [kl] and the reverse for [k#r] vs. [kɾ]. Therefore, clusters are produced with a more extreme lingual configuration across a word boundary than in syllable-onset position, which at least in part may be predicted by segmental factors for the [k#r]-[kɾ] pair. These articulatory data are compared with duration data for all sequence pairs.

Keywords: stop + liquid and liquid + stop clusters; Catalan; ultrasound

check for updates

Citation: Recasens, Daniel. 2024. The Effect of Manner of Articulation and Syllable Affiliation on Tongue Configuration for Catalan Stop–Liquid and Liquid–Stop Sequences: An Ultrasound Study. *Languages* 9: 233. https://doi.org/10.3390/languages9070233

Academic Editors: Rebeka Campos-Astorkiza and John Hajek

Received: 29 January 2024
Revised: 19 June 2024
Accepted: 21 June 2024
Published: 27 June 2024

1. Introduction

The present study is an ultrasound investigation of changes in tongue position in Catalan two-consonant sequences composed of an underlying stop and an alveolar lateral or rhotic as a function of differences in stop manner of articulation (voiced stop vs. voiced approximant) and syllable affiliation of the two consecutive consonants (heterosyllabic vs. tautosyllabic).

Regarding the first major research topic, articulatory differences as a function of the manner of articulation of the underlying stop consonant will be investigated through a comparison between the postvocalic sequences [kl] and [ɣl], and [kɾ] and [ɣɾ], in syllable-onset position, on the one hand, and between the heterosyllabic sequences [lk] and [lɣ], and [rk] and [rɣ], on the other hand. As to the presence of [ɣ] in the postvocalic sequences [ɣl] and [ɣɾ] and in the postconsonantal ones [lɣ] and [rɣ], it needs to be stated that, analogously to Spanish, Catalan has a phonological stop lenition rule according to which syllable-initial /b d g/ are realized as approximants after any non-stop segment with the exception of the sequences [ld ʎd] and to a large extent [fb] as well where C2 is realized as a stop because the two heterosyllabic consonants are homorganic (Recasens 1993). The consonant realizations [β ð ɣ] (which, in contrast with the International Phonetic Alphabet chart, are meant to be not fricative but approximant) are produced with non-turbulent airflow exiting the vocal tract through a relatively wide central constriction, which may occur at the velar zone (for [ɣ]), at the upper teeth ([ð]) or at the lips ([β]). However, as the studies for the Iberian languages mentioned below (see also Martínez-Celdrán 2004) and others for Tuscan Italian (Giannelli and Savoia 1978; Dalcher 2006) have shown, stop lenition is in fact a gradient process such that these approximants may differ a great deal in constriction opening degree especially when occurring postconsonantally, thus varying from more fricative-like to more vowel-like. Thus, it has been found for Spanish, Catalan and Galician in this respect that constriction narrowing for postconsonantal [β ð ɣ] is very much determined by the degree of constriction for the preceding consonant such that more open, less constricted approximant realizations are more prone to occur after sonorants than after fricatives and when the two consecutive consonants are heterorganic than when they are (quasi)-homorganic (Martínez-Celdrán and Regueira 2008; Hualde et al. 2011; Recasens 2015).

Through a comparison of the sequence pairs [lk]-[lɣ] and [rk]-[rɣ], the goal of this part of the study is to elicit whether some approximation of the tongue back to the pharyngeal wall or the velar region for contextual [l] and [r] may render the adjacent approximant [ɣ] fricative-like and therefore nearly as constricted as [k] or at least more constricted than postvocalic [ɣ] occurring in the sequences [ɣl] and [ɣɾ]. Several specific indications about the phonetic implementation of the lateral and the rhotic should be given at this stage. In Catalan, the alveolar lateral /l/ used to be dark, while at the present time darkness degree depends on the speaker and dialectal zone taken into consideration (Recasens 2014, pp. 175–204; Simonet 2010). While dark and clear /l/ share an apicoalveolar place of articulation, the two consonant varieties differ regarding tongue body configuration: clear /l/ is produced with a fronted tongue body and a raised blade and predorsum, which need to be somewhat braced laterally to allow for the passage of airflow through the mouth sides; dark /l/, on the other hand, exhibits a relatively lowered tongue front immediately behind the constriction location and a somewhat retracted postdorsum which approaches the velar and/or pharyngeal articulatory zones (Delattre 1965, p. 89; Ladefoged and Maddieson 1996, pp. 183–93). As to the alveolar rhotic, in Catalan, it is implemented as a tap in syllable-onset clusters and also morpheme-medial intervocalically (i.e., [ɾ] in the sequences [kɾ] and [ɣɾ], as in *acritud* 'acrimony' and *agrair* 'to express one's gratitude to', and also in *pare* 'father'), as a full trill absolute stem initially and after a heterosyllabic consonant (as in *rus* 'Russian', *pre-romà* 'pre-Roman', and also in *honra* 'honor') and as a short trill or as a tap in syllable-coda position (i.e., [r] in the sequences [rk] and [rɣ], as in *arca* 'arch' and *murga* 'lots of noise') (Recasens 1993, pp. 176–78). Regarding the articulatory characteristics of the two rhotic types, the alveolar trill is articulated with two or more fast apicoalveolar contacts (which may be reduced to a single long contact preconsonantally) and, analogously to dark /l/, a somewhat lowered tongue front and some tongue body retraction, while the production of the alveolar tap involves one fast ballistic apicoalveolar contact and a similar tongue body configuration to that for clear /l/ (see Recasens and Espinosa 2007 for Catalan and Proctor 2009 for the two rhotics in Spanish). Given a trend for the dark lateral and the alveolar trill to involve some tongue body retraction towards the velar and pharyngeal zones, it could be that the constriction degree for the velar approximant approaches that for the velar stop or at least that it is greater postconsonantally than postvocalically and therefore for [lɣ] and [rɣ] than for [ɣl] and [ɣɾ].

The first part of the study is also concerned with other aspects about the articulation of the velar approximant [ɣ]. In particular, we will explore the extent to which differences in constriction degree between the velar stop and the velar approximant cooccur with differences in tongue body position behind and in front of the velar zone. It may be that the formation of a wider constriction for [ɣ] than for [k] is assisted by shaping the tongue body back and the tongue predorsum in a certain way and that this consonant-specific lingual configuration serves to facilitate the passage of air throughout the vocal tract for the approximant. This research issue is all the more relevant since, to our knowledge, there are no data on overall tongue body position for [β ð ɣ] in the phonetics literature. On the other hand, we will look into whether there are changes in tongue position during the lateral and the rhotic, which depend on whether these consonants are flanked by the velar approximant or by the velar stop. Regarding the velar approximant, those consonant-to-consonant coarticulatory effects could seek to ease the production of [ɣ] by anticipating its overall tongue posture during the preceding liquid. Results from this research topic should improve our knowledge about the coarticulatory patterns which may occur in consonant sequences by showing that, not differently from stops, the syllable-initial approximants [β ð ɣ] in Catalan and Spanish are produced with active lingual or labial gestures which may overlap with the adjacent phonetic segments in the speech chain.

The present study will also look into whether sequences made of a velar stop C1 and a lateral or a rhotic C2 exhibit differences in lingual configuration depending on syllable affiliation and, therefore, on whether the two consonants occur in syllable-onset position or in different syllables. The motivation for this part of the study is to ascertain whether a syllable boundary effect operates in Catalan such that consonantal sequences which may occur syllable initially are produced less coherently in C#C sequences than word internally. The relationship between tongue configuration and syllable affiliation will be explored by comparing [k#l] to onset [kl] (as in *trec làmines* 'I take out metal plates' vs. *està clar* 'it is clear') and also [k#r] to onset [kɾ] (*sac rodó* 'round sack' vs. *ho creu* 'he/she believes it'). In Catalan, the sequences [k#l] and [k#r] are supposed to exhibit voicing throughout the velar stop due to the application of a regressive assimilation voicing rule, though in this particular case, the velar stop tends to be partly or fully voiceless (Recasens 2014, pp. 342–45). While substantial cross-linguistic research has been carried out on the articulatory implementation of the syllable-initial cluster [kl] (Gibbon et al. 1993; Pouplier et al. 2022), articulatory data about the ways in which this cluster differs from [k#l] have been collected mainly for English and German. Articulatory studies carried out on these Germanic languages reveal that prosodic boundary strength affects the degree of gestural overlap between consecutive consonants such that the stop and the lateral overlap less and lengthen more in the sequence [k#l] than in the onset cluster [kl] (Hardcastle 1985 and Byrd and Choi 2010 for English and Bombien et al. 2013 for German). These articulatory data are consistent with the claim made by Articulatory Phonology (Browman and Goldstein 1990, 1992) that an increase in gestural overlap should result in a decrease in gestural magnitude, thus implying that the stop ought to be longer, less overlapped, and more constricted in [k#l] than in [kl]. Acoustic data for other English consonant sequences also reveal that segmental duration is a systematic correlate of juncture, i.e., consonants lengthen when flanked by a word or a strong morphological boundary in comparison to when they are not (Lehiste 1960; Hoard 1966). While no data for this specific boundary effect appears to be available for either Catalan or other Romance languages, acoustic data for velar + /l/ sequences reported in Redford (2007) show that the effect in question is not universal: she found that the stop closure is enhanced across a word boundary ([k#l]) vs. syllable initially ([kl]) in English but less so or not at all in Finnish and Russian. Taking these data into consideration, the present study will look into whether the presence of a syllable/word boundary between two heterosyllabic consonants affects the articulatory manifestation and duration of at least the velar stop in Catalan such that it has a longer duration and is produced with a more extensive tongue dorsum gesture and thus a narrower dorsovelar constriction in [k#l] and [k#r] than in [kl] and [kɾ], respectively.

Another factor could play a role in the phonetic implementation of velar + rhotic clusters depending on whether occurring syllable initially or across a word boundary. The fact that the rhotic is realized as a tap in the onset cluster ([kɾ]) and as a full trill in the heterosyllabic sequence ([k#r]) could impinge on the articulatory configuration of the velar stop and on the characteristics of the C1-to-C2 lingual movement. In particular, in so far as the overall tongue body configuration for velars resembles that for [r] rather than that for [ɾ], one would expect less postdorsal movement and a more constricted velar stop in the sequence with the former rhotic ([k#r]) than in that with the latter ([kɾ]). If this is the case, both the syllable affiliation and the C2-to-C1 coarticulatory effects could lead to a similar tongue configuration during the velar consonant. It follows from this that the boundary effect under investigation should be ascertained in a more straightforward way in sequences with /l/ than in those with the rhotic in so far as in Catalan, the rhotic shows relevant articulatory differences associated with manner of articulation in the sequences [kɾ] and [k#r].

To recapitulate, the present investigation will look into differences in constriction degree and overall tongue configuration for [ɣ] as a function of contextual liquids sharing a similar tongue body configuration at the back of the vocal tract and into the extent to which the lingual gesture for the velar approximant is anticipated during the preceding liquid. The second major research topic deals with the syllable/word boundary effect on consonant articulation in the sequence pairs [kl]/[k#l] and [kɾ]/[k#r]. Results from this investigation should contribute to a better understanding of the articulatory characteristics and coarticulatory behavior of the velar approximant while throwing some light onto the production constraints on approximants of other places of articulation as well.

Ultrasound should allow for these issues to be looked into since it provides data on the configuration of the body of the tongue within the vocal tract and, consequently, about tongue postdorsum placement at the velar and pharyngeal regions. This recording and analysis technique has been previously used by ourselves in order to examine the tongue configuration characteristics of Catalan consonants (see, for example, Recasens and Rodríguez 2017), which puts us in a good position to interpret the articulatory data for the velar sequences with a liquid subject to analysis in the present investigation. In comparison to other techniques which are commonly used for the study of the articulation of consonants, such as electropalatography, electromagnetic articulography, and magnetic resonance (Kochetov 2020), an advantage of ultrasound is that it provides data on tongue postdorsum position at the upper and lower pharynx and at the velar zone and on tongue predorsum height at the palatal zone. A problem with this recording technique is that it barely provides data at the tongue tip and the tongue root since the mandible and hyoid bones refract the sound before it reaches the tongue surface, thus creating a black region where the tongue tip and the tongue root are located (see Stone 2005). This, however, should not be of great concern to the present investigation since much of the tongue configuration data of interest correspond to the palatal, velar, and upper pharyngeal regions of the vocal tract.

2. Materials and Methods

Tongue configuration and acoustic data were collected for the following pairs of consonant sequences: the syllable-onset clusters [kl]-[ɣl] and [kɾ]-[ɣɾ]; the heterosyllabic sequences [l#k]-[l#ɣ], [r#k]- [r#ɣ], and [k#l]-[k#r]. All sequences of consonants occurred in intervocalic position and were embedded in the meaningful Catalan sentences listed in Table 1. Sentences had four syllables and lexical stresses falling on the vowels preceding and following C1 and C2, respectively.

Table 1. List of Catalan sentences with English glosses. The consonant sequences under analysis appear underlined.

1. [kl]	un dinar <u>cl</u>au	'an important dinner'
2. [ɣl]	ell pujà <u>gl</u>aç	'he brought up ice'
3. [kɾ]	i menjà <u>cr</u>anc	'and (s)he ate crab'
4. [ɣɾ]	ell pujà <u>gr</u>ades	'he climbed steps'
5. [l#k]	anima<u>l c</u>ar	'expensive animal'
6. [l#ɣ]	va matar e<u>l g</u>all	'(s)he killed the rooster'
7. [r#k]	un rada<u>r c</u>ar	'an expensive radar'
8. [r#ɣ]	és un ba<u>r g</u>al	'it is a French bar'
9. [k#l]	és un fra<u>c l</u>àbil	'it is an outmoded tailcoat'
10. [k#r]	un ata<u>c r</u>àpid	'a fast attack'

Sentences were recorded six times by five native Catalan speakers, i.e., two men (DR, RO) and three women (ES, JU, IM) of 40–60 years of age who speak Catalan on a regular basis in their everyday lives. Ultrasound recordings were performed with an Echo Blaster unit type EB128CEXT from TELEMED and a microconvex Echo Blaster 128 CEXT transducer with a 2 to 4 MHz frequency range and a central curvature of 20 mm. The ultrasound images were acquired using a probe with a 100% of 104° field of view and a frequency of 2 MHz, which was attached to a transducer holder positioned under the subject's chin in an Articulate Instruments Stabilization Headset. The recording sampling rate was 54 frames per second, yielding one image every 18.5 ms. Image streams were recorded synchronously with the audio signal sampled at 22,050 Hz with an AKG-D70 microphone. Contours of the back of the alveolar zone and hard palate were also recorded by asking speakers to press the tongue against their hard palate. Tongue contours were tracked automatically at all temporal frames along each C#C sequence token for each speaker using the Articulate Assistant Advanced (AAA) software and adjusted manually by the paper author. Data points for all tongue contours were exported in ASCII files as x-y coordinates with their origin located at the bottom-left corner of the ultrasound image towards the rear of the vocal tract. Acoustic files were also exported in .wav format in order to extract segmental duration measures.

Segmentation was carried out on waveform and spectrographic displays. The boundary between C1 and C2 was taken to occur at the following events: in stop/approximant + liquid sequences, two temporal points were identified in case a C1 stop burst was present (i.e., C1 closure offset before the short stop burst, C2 onset immediately after the C1 stop burst), and a single temporal point at the beginning of the formant structure for the liquid in the absence of a burst; in liquid + stop/approximant sequences, we identified a single point at the offset of the C1 formant structure or else two temporal points at C1 offset and C2 onset whenever there was a vocalic portion at the offset of a rhotic C1. Depending on whether the boundary between C1 and C2 consisted of one or two temporal points, the lingual splines were processed at six points in time (at C1 onset, C1 midpoint, C1 offset, C2 onset, C2 midpoint, C2 offset) or at five points (at C1 onset, C1 midpoint, C1 offset/C2 onset, C2 midpoint, C2 offset). In the former case and as referred to above, the interval between C1 offset and C2 onset could be occupied by a stop burst or by a vocalic portion.

Tongue spline data points were converted from Cartesian to polar coordinates (Mielke 2015) by shifting the origin of the ultrasound image to approximately the center of the ultrasound probe, which was located at X = 86.7 mm and Y = 0 mm. SSANOVA smoothed splines consisting of strings of points separated by 0.01 radians with the associated standard errors were computed across the splines for all tokens of each consonant sequence using the R package gss to find a best-fit curve (Davidson 2006; Gu 2014). The rightmost and leftmost edges of the smoothed splines were determined by entering into the SSANOVA computation procedure the corresponding mean angle radian values across all tokens of the clusters of interest.

In order to elicit differences in tongue position at different regions of the tongue, the length of the SSANOVA splines displayed in Cartesian coordinates was divided into four portions which correspond to different articulatory zones, namely, alveolar (ALV), palatal (PAL), velar (VEL) and pharyngeal (PHAR), separately for each subject. This subdivision procedure was carried out by applying the same criterion as in the previous study Recasens and Rodríguez (2017) since the data for the clusters subject to analysis in the two studies were acquired in the same recording session. As described next, the criterion for determining the four articulatory zones involved the use of a different set of consonants and consonantal sequences from those subjected to analysis in this manuscript: the boundary between the alveolar and the dental zone was identified at an inflection point occurring at the spline front edge during dental /t/ in the sequence /pt/ and that between the alveolar and palatal zones at another inflection point located at the back alveolar area during the trill /r/ in the sequence /pr/ (/r/ is postalveolar in Catalan); the boundary between the palatal and velar zones was placed at the closure location for the velar stop in the sequence /iki/ which according to EPG data is articulated at the postpalatal zone, just in front of the soft palate, in Catalan; the length of the velar zone was taken to be 1.25 and 1.51 times that of the palatal zone in the case of the male and female speakers, respectively, as reported by Fitch and Giedd (1999); finally, the pharyngeal zone extended between the left edge of the velar zone all the way until the bottom edge of the lingual splines.

Distances between each of the four lingual regions and the origin of the ultrasound field of view were measured at all temporal points selected for analysis. The distance values at the velar (VEL) and palatal (PAL) zones were obtained by averaging the distances between the five central points at each zone and the origin. A different evaluation criterion was applied to the two extreme zones, alveolar (ALV) and pharyngeal (PHAR), in view of the fact that the splines for the consonant sequences subject to analysis could differ in length: in this case, we averaged the distances between the origin and five points located not at the zone midpoint but further away from the outer edges of the two zones in question, namely, at the upper third of the pharyngeal zone and at the leftmost third of the alveolar zone.

Linear Mixed Model (LMM) statistical tests were run on the distance values gathered at the midpoints of C1 and C2 with speaker as a random effect. Separate tests were performed on data for each of the cluster pairs of interest, i.e., [kl]-[ɣ̞l], [kɾ]-[ɣ̞ɾ], [lk]-[lɣ̞], [rk]-[rɣ̞], [k#l]-[kl], and [k#r]-[kɾ]. The LMM tests had the fixed effects ‘sequence’ (with 2 levels for each pair of consonant sequences, as for example [kl] and [ɣ̞l] in the case of the pair [kl]-[ɣ̞l]), ‘C place’ (with 2 levels corresponding to the velar and alveolar places of articulation) and ‘zone’ (with the 4 levels ALV, PAL, VEL, and PHAR). Additional LMM tests, one for each cluster pair, were also carried out on the cluster duration data using the duration values for all cluster tokens with ‘sequence’ and ‘C place’ as fixed effects. Least Significance Difference (LSD) post-hoc tests were run on all main effects and significant interactions in order to find out whether numerical differences between pairs of levels of a given statistical variable reached significance or not. Given the large number of tests involved in the LMM analyses, the Benjamini–Hochberg (BH) correction procedure for adjusting the false discovery rate was applied to those variable comparisons that were of relevance to the present investigation. In all statistical tests, the significance level was set at $p < 0.05$.

3. Results

Statistical results for the main effects and factor interactions obtained from the LMM tests run on the distance and duration values are given for all pairs of clusters in Tables 2–5. These statistical results will be commented on with reference to the lingual position and duration data and to the tongue profiles displayed in Figures 1–6.

Statistical differences among distance values for different articulatory zones, and thus for the main zone effect and for the C place × zone interaction, will not be reported. It suffices to say in this respect that distances between the tongue and the origin of the ultrasound field of view were consistently larger at the velar (VEL) and palatal (PAL) zones than at the more extreme pharyngeal (PHAR) and alveolar (ALV) zones. Moreover, they were larger at the velar zone for velars than for alveolars and the reverse at the front of the vocal tract, which is in agreement with differences in place of articulation between the two consonant classes.

3.1. Effect of Manner of Articulation

3.1.1. Lingual Configuration

Table 2 shows the statistical results for the distance values between the tongue and the origin of the ultrasound field of view in the case of clusters with the velar stop and the velar approximant. They indicate a significant main effect of sequence and C place of articulation for the [rk]-[rɣ̞] pair, which is associated with larger distances for [rk] vs. [rɣ̞] and for the velar vs. the rhotic. There are two relevant significant interactions, i.e., sequence × zone and sequence × zone × C place. According to the former interaction, tongue distances are significantly greater for sequences with a voiced approximant than for those with a voiceless stop at PHAR ([kɾ]-[ɣɾ], [lk]-[lɣ]) and the reverse at PAL ([kɾ]-[ɣɾ], [lk]-[lɣ], [rk]-[rɣ]). The triple sequence × zone × C place interaction, which achieves significance only for [lk]-[lɣ], is associated with larger distances between [k] and [ɣ] at PHAR and PAL when compared to the two other articulatory zones.

Data for the distance values at successive points in time for the four pairs of sequences plotted in Figure 1 reveal indeed the presence of larger distances at the pharynx for the sequences with the voiced approximant (discontinuous lines) than for those with the voiceless stop (continuous lines), and the opposite relationship at the velar and palatal zones. They also show that, while differences in tongue position between the two sequence pairs occur mainly during the velar consonant (at C1 in the case of [kl]-[ɣl] and [kɾ]-[ɣɾ], at C2 in the case of [lk]-[lɣ] and [rk]-[rɣ]), they may also extend into the adjacent liquid. This coarticulatory effect applies mostly to the coda consonants [l] and [r], which tend to be either more retracted at PHAR and thus at the back of the vocal tract (mostly [l]) or lower at PAL and thus at the palatal zone (mostly [r]) whenever occurring before [ɣ] than before [k].

Table 2. F and *p* values for the main effects and factor interactions obtained from LMM analyses run on the distance values between the tongue surface and the origin of the ultrasound field of view for cluster pairs with a velar stop and a velar approximant (*, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$). NS: non-significant effect.

	kl-ɣl	kɾ-ɣɾ	lk-lɣ	rk-rɣ
Sequence	NS	NS	NS	4.86 (1, 52) *
C place	NS	NS	NS	65.13 (1, 52) ***
Zone	8.93 (3, 54) **	9.02 (3, 44) **	9.33 (3, 48) **	6.80 (3, 52) **
Sequence × zone	NS	6.63 (3, 44) ***	11.19 (3, 48) ***	3.50 (3, 52) *
C place × zone	12.83 (3, 54) ***	70.67 (3, 44) ***	22.56 (3, 48) ***	29.55 (3, 52) ***
Sequence × C place	NS	NS	NS	NS
Sequence × zone × C place	NS	NS	4.53 (3, 48) **	NS

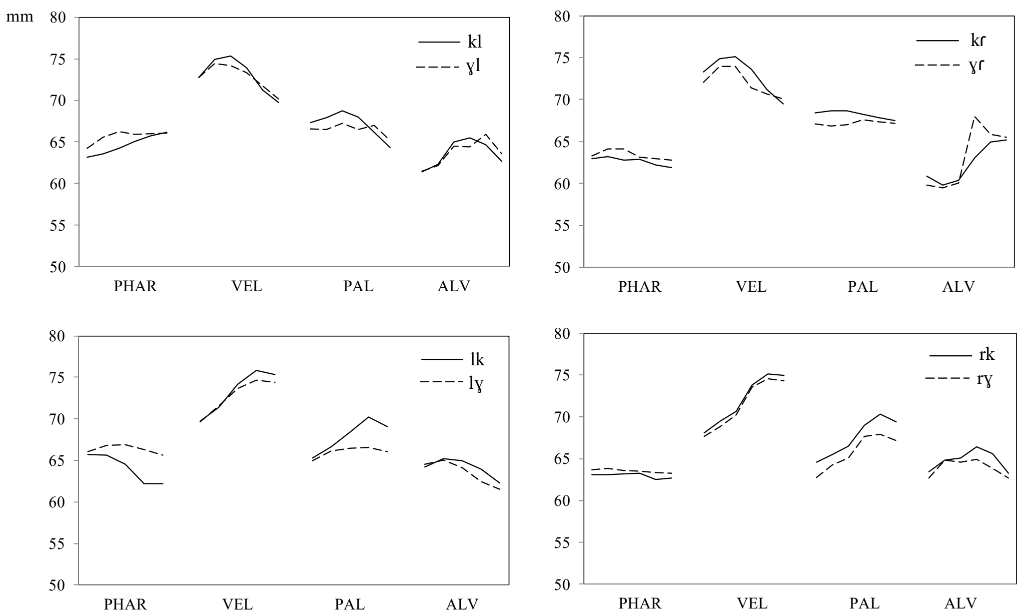

Figure 1. Cross-speaker distance measures between the tongue surface and the origin of the ultrasound field of view sampled at consecutive temporal points during the sequence pairs [kl]-[ɣl], [kɾ]-[ɣɾ], [lk]-[lɣ], and [rk]-[rɣ]. The distance trajectories for each articulatory zone proceed from C1 onset (leftmost edge) to C2 offset (rightmost edge) through intermediate temporal points. PHAR = pharyngeal, VEL = velar, PAL = palatal, ALV = alveolar.

The distance data for [kl]-[ɣl] and [kɾ]-[ɣɾ] plotted in Figure 1 are in agreement with the lingual configurations represented in Figure 2 for speaker JU. Figure 2 reveals indeed that, in comparison with C1 = [k], C1 = [ɣ] tends to be produced with a lower tongue dorsum position and thus a wider dorsal constriction at the palatal zone and a more retracted postdorsum at the pharynx.

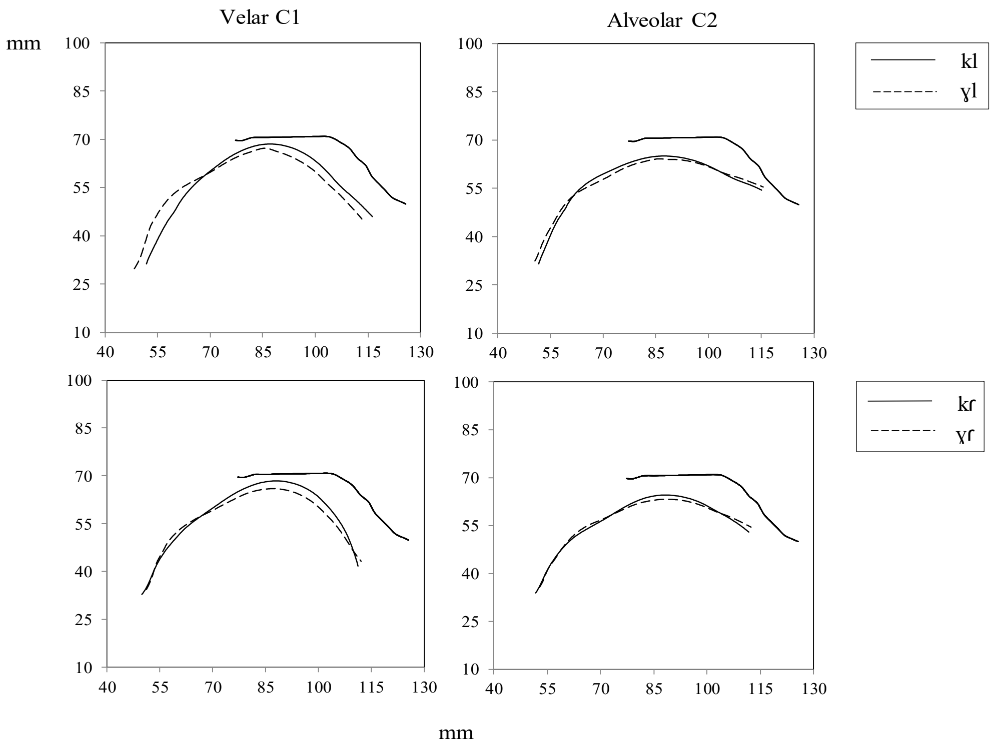

Figure 2. Tongue configurations at the midpoint of C1 and C2 for [kl] and [ɣl] (**top**) and [kɾ] and [ɣɾ] (**bottom**) with the palate trace superimposed according to speaker JU. The front of the mouth is on the right of the graphs.

3.1.2. Duration

Figure 3 reports duration data for the two consonants of all consonant sequences subjected to analysis in this part of the study. According to the corresponding statistical results provided in Table 3, the clusters with a rhotic show a main sequence effect and a main C place effect, which happen to be associated, respectively, with a longer duration for [kɾ] vs. [ɣɾ] and for [rk] vs. [rɣ] and with a shorter duration for the rhotic than for the velar consonant whether in onset or coda position. There is also a significant sequence × C place interaction for the [rk]-[rɣ] sequence pair, which turned out to be related to differences in duration for velar > rhotic and for [k] > [ɣ] and a somewhat longer rhotic before [ɣ] than before [k]. Figure 3 also shows that the alveolar tap is longer after [k] than after [ɣ] in the sequence pair [kɾ]-[ɣɾ], though this difference did not achieve significance. As to the clusters with the alveolar lateral, there is only a significant sequence × C place interaction for [kl]-[ɣl], which is associated with a longer velar consonant than alveolar liquid in the case of [kl] and the reverse for [ɣl], and a longer velar for [kl] vs. [ɣl] and a longer alveolar lateral for [ɣl] vs. [kl].

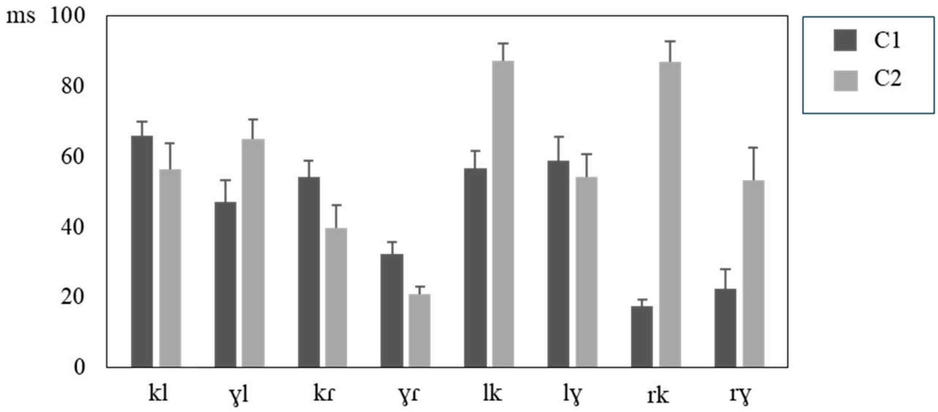

Figure 3. Consonant durations and standard deviation values for cluster pairs differing in manner of articulation for the velar consonant.

Table 3. F and *p* values for the main effects and factor interaction obtained from LMM analyses run on the duration values for cluster pairs with a velar stop and a velar approximant (*, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$). NS: non-significant effect.

	kl-ɣl	kɾ-ɣɾ	lk-lɣ	rk-rɣ
Sequence	NS	50.84 (1, 100) **	NS	12.59 (1, 100) *
C place	NS	16.96 (1, 100) *	NS	149.5 (1, 100) ***
Sequence × C place	14.7 (1, 100) *	NS	NS	10.16 (1, 100) *

A longer rhotic before [ɣ] than before [k] and a longer lateral after [ɣ] than after [k] indicate that the C1-to-C2 closing movement for [rk] and [rɣ] and the C1-to-C2 opening movement for [kl] and [ɣl] proceed more slowly in sequences with the approximant [ɣ] than in those with the stop [k].

3.2. *Syllable Affiliation Effect*

3.2.1. Lingual Configuration

Statistical results for the distance data between the tongue surface and the origin of the ultrasound field of view for the heterosyllabic stop + liquid sequences are given in Table 4. Results from the post-hoc test run on the main sequence effect reveal significantly greater distances across zones for [k#l] than for [kl] but no differences between [k#r] and [kɾ], and those performed on the main C place effect, larger distances for the velar than for the liquids. Moreover, a significant sequence × zone interaction turned out to be associated with a more distant tongue position from the origin for [k#r] than for [kɾ] at ALV and PHAR and the reverse at PAL. These sequence- and consonant-dependent differences may be seen in the tongue distance trajectories over time plotted in Figure 4 and are also apparent in the lingual splines for speaker RO displayed in Figure 5. There is then a trend for the voiceless velar stop to be produced with a more extreme articulation in heterosyllabic than in tautosyllabic sequences. As suggested in the Introduction section, in sequences with the alveolar rhotic, this difference in tongue configuration could be associated not only with the presence vs. absence of a syllable juncture but also with differences in C2-to-C1 coarticulation in so far as the alveolar trill is articulated with a somewhat lower front dorsum and a more retracted tongue body than the alveolar tap.

Table 4. F and *p* values for the main effects and factor interactions obtained from LMM analyses run on the distance values between the tongue surface and the origin of the ultrasound field of view for cluster pairs differing in word boundary availability (*, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$). NS: non-significant effect.

	k#l-kl	k#r-kɾ
Sequence	6.82 (1, 52) *	NS
C place	6.16 (1, 52) *	7.39 (1, 50) **
Zone	7.84 (3, 52) **	9.16 (3, 50) **
Sequence × zone	NS	5.79 (3, 50) **
C place × zone	11.94 (3, 52) ***	32.66 (3, 50) ***
Sequence × C place	NS	NS
Sequence × zone × C place	NS	NS

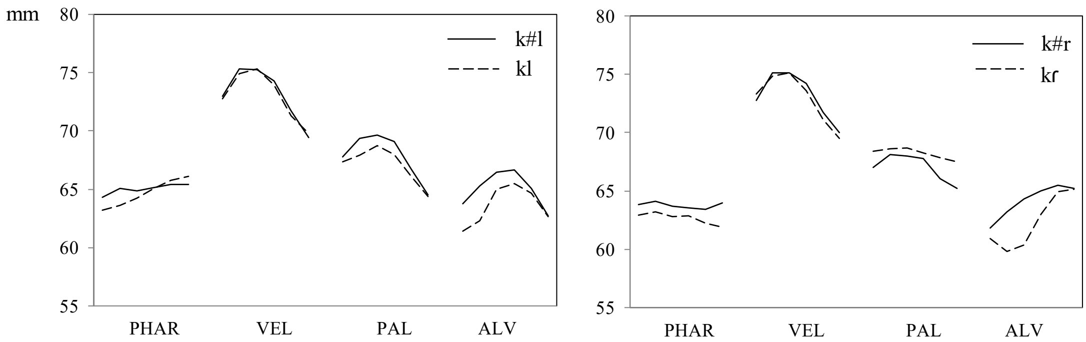

Figure 4. Cross-speaker distance measures between the tongue surface and the origin of the ultrasound field of view sampled at consecutive temporal points during the sequence pairs [k#l]-[kl] and [k#r]-[kɾ]. The distance trajectories for each articulatory zone proceed from C1 onset (leftmost edge) to C2 offset (rightmost edge) through intermediate temporal points. PHAR = pharyngeal, VEL = velar, PAL = palatal, ALV = alveolar.

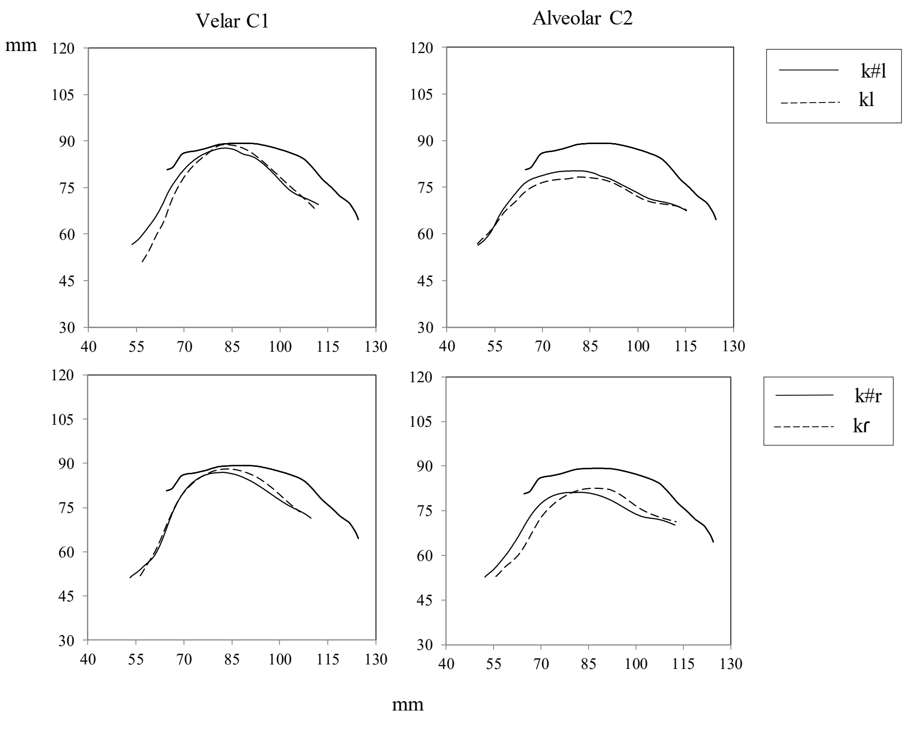

Figure 5. Tongue configurations at the midpoint of C1 and C2 for [k#l] and [kl] (**top**) and [k#r] and [kɾ] (**bottom**) with the palate trace superimposed according to speaker RO. The front of the mouth is on the right of the graphs.

3.2.2. Duration

Mean duration values for the cross-boundary and syllable-onset consonant sequences and the corresponding statistical results are provided in Figure 6 and Table 5, respectively. According to the data reported in the figure and the table, while [k#l] and [kl] do not differ significantly from each other, there is a significantly greater duration for [k#r] than for [kɾ] (main sequence effect), which is associated with a much longer realization for the trill [r] than for the tap [ɾ] (sequence × C place interaction).

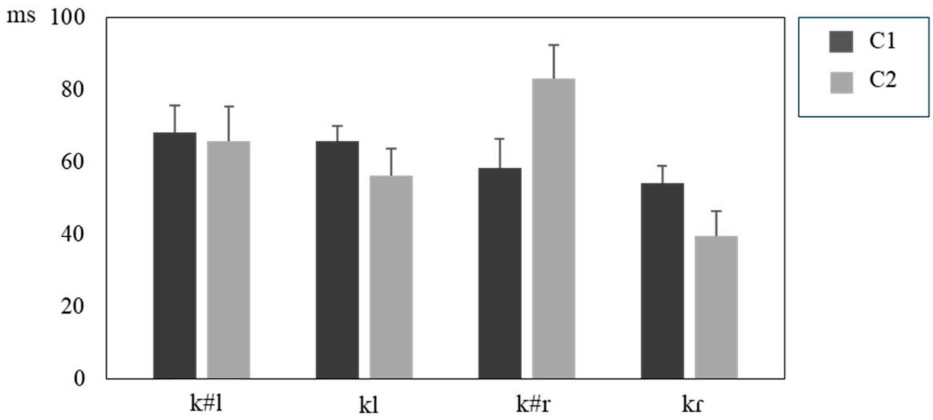

Figure 6. Consonant durations and standard deviation values for cluster pairs differing in word boundary availability.

Table 5. F and *p* values for the main effects and factor interaction obtained from LMM analyses run on the duration values for cluster pairs differing in word boundary availability (*, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$). NS: non-significant effect.

	k#l-kl	k#r-kɾ
Sequence	NS	15.32 (1, 100) *
C place	NS	NS
Sequence × C place	NS	10.77 (1, 100) *

4. Discussion

Articulatory data evaluated in terms of spatial distance for the cluster pairs [kl]-[ɣl], [kɾ]-[ɣɾ], [lk]-[lɣ], and [rk]-[rɣ] show that the (shorter) velar approximant is less constricted than the (longer) voiceless velar stop at the velar/palatal zone while exhibiting a more retracted tongue body at the pharynx as a general rule. Therefore, it appears that the loss of central contact at the velar constriction location (and also at the palatal zone where the predorsum is located) for the approximant as compared to the voiceless stop occurs concomitantly with some postdorsum backing. This finding about tongue position behind and in front of the constriction location for the velar approximant, which has not been reported previously in the phonetics literature, may result from the formation of a wide velar constriction while also being associated with the need to allow continuous airflow through the oral cavity. It is important to emphasize that differences in overall tongue body configuration between [k] and [ɣ] are not the same as those between [k] and [g], which should be associated with the voicing contrast between the two stops but not with a difference in manner of articulation. In comparison to voiceless stops, voiced stops of the same place of articulation in English and other languages exhibit an active expansion of the supraglottal cavity system, which is achieved mainly by fronting the tongue body and thus enlarging the pharyngeal cavity and, albeit less consistently, by lowering the tongue front and thus enlarging the oral cavity (Westbury 1983; Ahn 2018). These changes in lingual configuration facilitate the maintenance of voicing by lowering the intraoral pressure level above the glottis. This piece of evidence confirms that a more retracted tongue body (and, to a large extent, a lower tongue predorsum position) for [ɣ] with respect to [k] depends not on differences in voicing but on differences in the manner of articulation between the two consonantal realizations.

Visual inspection of the distance trajectories reported in Figure 1 reveals essentially the same tongue position for [ɣ] after a vowel and after the lateral and the rhotic. Thus, while the lateral and the rhotic may involve some tongue body retraction, this retraction movement does not appear to be prominent enough to cause the constriction narrowing degree for [ɣ] to approach that for [k], which does not match with the hypothesis set in the Introduction that there should be differences in dorsovelar constriction degree between postvocalic and postconsonantal [ɣ]. This finding may be related to the fact that there is no central closure at the velar zone for [l] or [r], which could block the passage of air through the dorsal constriction for [ɣ]. Moreover, the fact that constriction degree is essentially the same postvocalically as postconsonantally may also be related to the difficulty involved in forming a complete dorsal closure at the soft palate for true velars, which may result in air leakage through the dorsal constriction, most especially when /g/ is preceded by a non-stop consonant (Kingston 2008, p. 21). Therefore, it turns out that /g/ is likely to be realized either as a stop or as a quite unconstricted approximant, much independent of the degree of opening for the preceding phonetic segment. In line with this possibility, previous studies have reported a greater lenition degree for /g/ than for /b/ and /d/ (see, for example, Dalcher 2006 for Florentine Italian and Tang et al. 2023 for Spanish).

Manner of articulation-dependent differences in tongue configuration between the stop and the approximant were found to extend into the preceding liquid in clusters with a coda lateral or rhotic, whether at the back of the vocal tract or at the palatal zone. Moreover, [k] turned out to be longer than [ɣ], as expected, while the liquid was longer when flanked by [ɣ] than by [k], which may be attributed to slower rates of articulatory movement as the lateral moves towards and out of an approximant vs. a stop. Evidence for a much slower rate of increase of the constriction area for approximants than for stops has been reported in the literature (Stevens 1999, p. 532). These findings provide some support for the implementation of syllable-initial [β ð ɣ] in Catalan and Spanish by means of labial or lingual gestures, which may be activated during the preceding phonetic segment in the speech chain.

In agreement with English and German data provided in the Introduction section (Byrd and Choi 2010; Bombien et al. 2013), the Catalan voiceless velar stop + liquid sequences were produced with a more extreme lingual configuration and thus less gestural overlap whenever occurring across a word boundary than syllable initially. This more prominent lingual configuration is implemented differently in sequences with an alveolar lateral and in those with an alveolar rhotic due presumably to the articulatory requirements involved in the production of the two liquids. Indeed, in stop + liquid sequences across a word boundary, the tongue body for [k] is more retracted at the pharynx (PHAR) and may be higher at the palatal zone (PAL) whenever C2 is [l], and also more retracted at PHAR but lower at PAL when C2 is the rhotic. Even though the ultrasound data does not provide information about the degree of dorsal contact at the velar place of articulation, differences in tongue configuration between [k#l] and [kl], i.e., a higher predorsum and a more advanced postdorsum in the former sequence than in the latter, suggest that the velar stop is produced with a greater dorsovelar closure area across a word boundary than word internally. As to the consonantal sequences with the rhotic, rhotic type may account for a lower predorsum position and a more retracted postdorsum for C1 in the case of [k#r] vs. [kɾ] since the alveolar trill is expected to be articulated with less predorsum height and more postdorsum backing than the alveolar tap (see Section 1).

In contrast with data for the Germanic languages showing duration differences for [k#l] > [kl], in Catalan, differences in articulatory displacement between velar stop + liquid sequences with and without a syllable boundary turned out not to be clearly correlated with the corresponding duration data in so far as [k#l] was not statistically longer than [kl] while duration differences between [k#r] and [kɾ] could be due to rhotic type rather than to syllable affiliation. The finding that /k#l/ differs from /kl/ in articulation but not in duration suggests that when produced in a VC#CV string, consonantal sequences that may occur syllable initially are produced more coherently and thus overlap to a larger extent in Catalan (and perhaps in other Romance languages as well) than in English or German even though the syllabification pattern is essentially the same, i.e., (VC)(CV), in all cases. This finding, which awaits further investigation, could be associated with a stronger trend for phonetic segments to lengthen and shorten and thus to be less isochronous, depending on the position that they occupy within the foot unit or the interstress interval in stress-timed languages like German and English than in syllable-timed languages like Catalan (Dauer 1983; Roseano et al. 2022).

The findings of the present investigation may be summarized as follows. The first part of the paper brings about new data for Catalan on the articulation of the velar approximant, mostly regarding its overall lingual configuration characteristics. The ultrasound data reveal that [ɣ] is implemented through a specific tongue configuration, which could serve an aerodynamic goal and small changes in constriction width as a function of context, which may follow from the difficulty involved in forming a complete dorsal closure at the soft palate. Moreover, coarticulatory effects in segmental articulation and duration suggest that [ɣ] is produced with an active articulatory gesture, which is prepared ahead of time during the preceding liquid. As to the effect of a syllable/word boundary on sequences composed of /k/ and a following liquid, the velar stop turned out to exhibit a more extreme lingual configuration but not a longer closure when flanked by a syllable/word boundary, a finding which in the case of rhotic clusters goes hand in hand with differences in manner of articulation for the rhotic. All in all, it is hypothesized that the boundary effect in question may be more straightforward in English and German than in Catalan and perhaps other Romance languages and that this could be related to differences in rhythmic organization between stress-timed and syllable-timed languages. Future research could explore context-dependent differences in articulatory implementation (both in constriction opening and in overall tongue body configuration) for stops vs. approximants of labial, dental, and velar places of articulation and differences in the boundary effect on articulation for velar + liquid clusters among different language types.

Funding: This research was funded by project PGC2018-096877-B-I00 of the Spanish Ministry of Science, Innovation and Universities and by the research group 2017 SGR 34 of the Catalan Government.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Additional data supporting the experimental results reported in the present study are unavailable.

Acknowledgments: I would like to thank Clara Rodríguez for her help with the data processing and statistical analysis and three reviewers for their insightful comments on a previous version of the manuscript.

Conflicts of Interest: The author declares no conflicts of interest.

References

Ahn, Suzy. 2018. The role of tongue position in laryngeal contrasts: An ultrasound study of English and Brazilian Portuguese. *Journal of Phonetics* 71: 451–67. [CrossRef]

Bombien, Lasse, Christine Mooshammer, and Philip Hoole. 2013. Articulatory coordination in word-initial clusters of German. *Journal of Phonetics* 41: 546–61. [CrossRef]

Browman, Catherine P., and Louis Goldstein. 1990. Tiers in articulatory phonology, with some implications for casual speech. In *Papers in Laboratory Phonology I: Between the Grammar and Physics of Speech*. Edited by John Kingston and Mary E. Beckman. Cambridge: Cambridge University Press, pp. 341–76.

Browman, Catherine P., and Louis Goldstein. 1992. Articulatory phonology: An overview. *Phonetica* 49: 222–34. [CrossRef]

Byrd, Dani, and Susie Choi. 2010. At the juncture of prosody, phonology, and phonetics. The interaction of phrasal and syllable structure in shaping the timing of consonant gestures. In *Laboratory Phonology 10*. Edited by Cécile Fougeron, Barbara Kühnert, Mariapaola D'Imperio and Nathalie Vallée. Berlin and New York: De Gruyter, pp. 31–59.

Dalcher, Christina V. 2006. Consonant Weakening in Florentine Italian: An Acoustic Study of Gradient and Variable Sound Change. Ph.D. dissertation, Georgetown University, Washington, DC, USA.

Dauer, Rebecca M. 1983. Stress-timing and syllable-timing reanalyzed. *Journal of Phonetics* 11: 51–62. [CrossRef]

Davidson, Lisa. 2006. Comparing tongue shapes from ultrasound imaging using smoothing spline analysis of variance. *Journal of the Acoustical Society of America* 120: 407–15. [CrossRef] [PubMed]

Delattre, Pierre. 1965. *Comparing the Phonetic Features of English, French, German and Spanish: An Interim Report*. Heidelberg: Julius Gross Verlag.

Fitch, W. Tecumseh, and Jay Giedd. 1999. Morphology and development of the human vocal tract: A study using magnetic resonance imaging. *Journal of the Acoustical Society of America* 106: 1511–22. [CrossRef] [PubMed]

Giannelli, Luciano, and Leonardo M. Savoia. 1978. L'indebolimento consonantico in Toscana I. *Rivista Italiana di Dialettologia* 2: 25–58.

Gibbon, Fiona, William Hardcastle, and Katerina Nicolaidis. 1993. Temporal and spatial aspects of lingual coarticulation in /kl/ sequences: A cross-linguistic investigation. *Language and Speech* 36: 261–77. [CrossRef] [PubMed]

Gu, Chong. 2014. Smoothing spline ANOVA models: R Package gss. *Journal of Statistical Software* 58: 1–25. [CrossRef]

Hardcastle, William J. 1985. Some phonetic and syntactic constraints on lingual coarticulation during /kl/ sequences. *Speech Communication* 4: 247–63. [CrossRef]

Hoard, James E. 1966. Juncture and syllable structure in English. *Phonetica* 15: 96–109. [CrossRef]

Hualde, José Ignacio, Ryan Shosted, and Daniel Scarpace. 2011. Acoustics and articulation of Spanish /d/ spirantization. Paper presented at XVII International Congress of Phonetic Sciences, Hong Kong, China, August 17–21; Edited by Wai Sum Lee and Eric Zee. Hong Kong: City University of Hong Kong, pp. 906–9.

Kingston, John. 2008. Lenition. In *Selected Proceedings of the 3rd Conference on Laboratory Approaches to Spanish Phonology*. Edited by Laura Colantoni and Jeffrey Steele. Somerville: Cascadilla Press, pp. 1–31.

Kochetov, Alexei. 2020. Research methods in articulatory phonetics I: Introduction and studying oral gestures. *Language and Linguistics Compass* 14: e12368. [CrossRef]

Ladefoged, Peter, and Ian Maddieson. 1996. *The Sounds of the World's Languages*. Oxford: Blackwell.

Lehiste, Ilse. 1960. An acoustic-phonetic study of internal open juncture. *Phonetica* 5 S1: 5–54. [CrossRef]

Martínez-Celdrán, Eugenio. 2004. Problems in the classification of approximants. *Journal of the International Phonetic Association* 34: 201–9. [CrossRef]

Martínez-Celdrán, Eugenio, and Xosé L. Regueira. 2008. Spirant approximants in Galician. *Journal of the International Phonetic Association* 38: 51–68. [CrossRef]

Mielke, Jeff. 2015. An ultrasound study of Canadian French rhotic vowels with polar smoothing spline comparisons. *Journal of the Acoustical Society of America* 137: 2858–69. [CrossRef] [PubMed]

Pouplier, Marianne, Manfred Pastätter, Philip Hoole, Stefania Marin, Ioana Chitoran, Tomas Lentz, and Alexei Kochetov. 2022. Language and cluster-specific effects in the timing of onset consonant sequences in seven languages. *Journal of Phonetics* 93: 1–28. [CrossRef]

Proctor, Michael I. 2009. Gestural Characterization of a Phonological Class: The Liquids. Ph.D. dissertation, Yale University, New Haven, CT, USA.

Recasens, Daniel. 1993. *Fonètica i fonologia*. Barcelona: Enciclopèdia Catalana.

Recasens, Daniel. 2014. *Fonètica i fonologia experimentals del català*. Barcelona: Institut d'Estudis Catalans.

Recasens, Daniel. 2015. The effect of contextual consonants on voiced stop lenition: Evidence from Catalan. *Language and Speech* 59: 1–23. [CrossRef]

Recasens, Daniel, and Aina Espinosa. 2007. Phonetic typology and positional allophones for alveolar rhotics in Catalan. *Phonetica* 63: 1–28. [CrossRef] [PubMed]

Recasens, Daniel, and Clara Rodríguez. 2017. Lingual articulation and coarticulation for Catalan consonants and vowels: An ultrasound study. *Phonetica* 74: 125–56. [CrossRef]

Redford, Melissa. 2007. Word-internal versus word-peripheral consonantal duration patterns in three languages. *Journal of the Acoustical Society of America* 121: 1665–78. [CrossRef] [PubMed]

Roseano, Paolo, Patrícia Marsà Morales, and Laura Alañá Vilas. 2022. Is Catalan syllable timed? An answer based on cluster analysis. In *Propuestas en fonética experimental: Enfoques metodológicos y nuevas tecnologías*. Edited by Beatriz Blecua, Jordi Cicres, Marina Espejel and María Machuca. Girona: Universitat de Girona, pp. 270–74.

Simonet, Miquel. 2010. Dark and clear laterals in Catalan and Spanish: Interaction of phonetic categories in early bilinguals. *Journal of Phonetics* 38: 663–78. [CrossRef]

Stevens, Kenneth N. 1999. *Acoustic Phonetics*. Boston: The MIT Press.

Stone, Maureen. 2005. A guide to analysing tongue motion from ultrasound images. *Clinical Linguistics and Phonetics* 19: 455–501. [CrossRef] [PubMed]

Tang, Kevin, Ratree Wayland, Fenqi Wang, Sophia Vellozzi, Rahul Sengupta, and Lori Altmann. 2023. From sonority hierarchy to posterior probability as a measure: The case of Spanish stops. *Journal of the Acoustical Society of America* 153: 1191–203. [CrossRef]

Westbury, John R. 1983. Enlargement of the supraglottal cavity and its relation to stop consonant voicing. *Journal of the Acoustical Society of America* 73: 1322–36. [CrossRef]