Research Article

A study on coarticulatory resistance and aggressiveness for front lingual consonants and vowels using ultrasound

Daniel Recasens a,b,*, Clara Rodríguez a,b

a Departament de Filològia Catalana, Universitat Autònoma de Barcelona, 08193 Bellaterra, Barcelona, Spain
b Institut d’Estudis Catalans, c/Carme, 47, Barcelona 08001, Spain

ARTICLE INFO

Received 21 December 2015
Received in revised form 18 September 2016
Accepted 22 September 2016

ABSTRACT

A new method for quantifying contextual variability at different regions of the tongue using ultrasound spline data reveals that tongue body coarticulatory resistance for Catalan consonants and vowels in VCV sequences decreases in the progression [ʎ, ɲ, ʃ] > [s, r] > [l, ɾ, t, n] > [ð] and [i, e] > [a] > [o] > [u]. These consonant and vowel hierarchies support the degree of articulatory constraint model of coarticulation according to which coarticulatory resistance depends on whether a given lingual region is involved in the formation of a closure or constriction and on the severity of the manner of articulation requirements. Data show that this coarticulatory scenario holds not only at the palatal zone, as revealed by previous coarticulation studies, but at the velar and pharyngeal zones as well. Partial exceptions are [s] and [i], which may allow for some more contextual variability than expected at the back of the vocal tract. Another major finding is that tongue body coarticulatory resistance and aggressiveness are highly positively correlated. The implications of these experimental results for speech production organization and sound change are discussed.

1. Introduction

Attention has been paid to the notions ‘coarticulatory resistance’ and ‘coarticulatory aggressiveness’ (also referred to as ‘coarticulatory dominance’) in the phonetics literature, namely, to whether consonants and vowels allow more or less coarticulation from adjacent segments, and whether those phonetic segments that are more resistant exert stronger coarticulatory effects on neighbouring phonetic segments than those that are less resistant (Fowler & Saltzman, 1993, Iskarous et al., 2013). Coarticulatory resistance is inversely related to coarticulatory sensitivity, i.e., a given phonetic segment becomes resistant the more it blocks changes in articulatory configuration exerted by the contextual phonetic segments in speech. For example, data on tongue dorsum coarticulation reveal that, in comparison to the alveolar nasal stop [n], the palatal nasal [ɲ] is at the same time more resistant to the influence of the contextual vowels and more aggressive regarding the consonant-to-vowel effects (Recasens & Espinosa, 2009). In particular, differences in tongue dorsum height as a function of [i] vs [a] are greater during [n] than during [ɲ] mostly because [i] causes tongue dorsum raising to take place during the alveolar consonant while [a] causes little tongue dorsum lowering to occur during the palatal consonant. On the other hand, effects from [n] and [ɲ] during adjacent [i] are small while those on the low vowel [a] are larger for the former consonant than for the latter and involve tongue dorsum raising. Regarding vowels, high front [i] happens to be more resistant than [a, u] to tongue body coarticulatory effects from consonants (see above), while causing greater changes in tongue body position on consonants like [ɲ] (Chen, Chang, & Iskarous, 2015).

According to the degree of articulatory constraint model of coarticulation (DAC), coarticulatory resistance and aggressiveness for a given articulatory structure depend directly on whether this articulator takes part in constriction or closure formation and also on the specific manner of articulation demands for the target segment (Recasens, Pallarès, & Fontdevila, 1997). Thus, as revealed by the studies just cited, palatal consonants are more constrained than several dentoalveolars since the involvement of the tongue dorsum (i.e., the superior surface of the tongue behind the blade) should constrain to a large extent the tongue body (i.e., the whole tongue from tip to root), while the formation of an apical or laminal closure or constriction leaves the tongue dorsum, postdorsum and root

* Corresponding author.
E-mail addresses: daniel.recasens@uab.es (D. Recasens), clara.isabel.rodriguez@uab.cat (C. Rodríguez).
Regarding place of articulation, coarticulatory resistance and aggressiveness ought to be higher for palatal consonants than to coarticulatory dominance (Recasens & Espinosa, 2009). The articulatory analysis of some additional front lingual consonants and whether an increase in degree of constraint for [s, r] renders these consonants equally resistant as [l, r] and the approximant [o, u]. Several aspects of the articulation of these linguistic sounds need to be pointed out. In Catalan, the stop consonant [t] is dentoalveolar control on constriction degree, approximants such as [l, r] are palatoalveolar fricative and [ɾ, t, n, s, r] are apical, and [ɾ, t, n, s] are laminopredorsal.

A good reason for using ultrasound data for the analysis of lingual coarticulation is in order to verify whether the predicted differences in coarticulatory resistance and aggressiveness among consonants and vowels hold not only at the tongue front and dorsal and thus at the alveolar and palatal zones, as shown by previous electropalatographic (EPG) and electromagnetic articulometer (EMA) studies, but also at the tongue posterior and root and thus at the velar and pharyngeal zones. Another advantage of ultrasound in comparison to EPG and EMA is that it provides information about tongue position at a large number of flesh points and therefore allows us to measure coarticulatory patterns more precisely.

A large scale ultrasound study on lingual coarticulation for several consonants and vowels such as the present one was lacking in the phonetics literature (previous ultrasound studies such as Zharkova and Hewlett (2009), Zharkova, Hewlett, and Hardcastle, (2011, 2012) and Noiray, Ménard, and Iskarous (2013) have dealt only with lingual fricatives and less so stops). This paper investigates the patterns of lingual coarticulatory resistance and aggressiveness for a more complete set of Catalan front lingual consonants and vowels than in previous EMA and EPG studies, where only [l, r, t, n, s, ñ] and [l, a, u] were subject to investigation and more attention was paid to coarticulatory resistance than to coarticulatory dominance (Recasens & Espinosa, 2009). The articulatory analysis of some additional front lingual consonants and vowels, i.e., [l, n, s] are generally apicalominal, [ñ] and [l, r, r] are apical, and [ñ, n, ñ] are laminopredorsal.

Given the place, manner and articulatory constraint characteristics of the Catalan front consonants and vowels described above, several predictions may be formulated about the degree of coarticulatory resistance and aggressiveness that these phonetic segments ought to show in symmetrical VCV sequences:

(a) Regarding place of articulation, coarticulatory resistance and aggressiveness ought to be higher for palatal consonants articulated with the tongue predorsum or with the blade and predorsum ([ʃ, n, ñ]) than for dentoalveolars produced with the tongue tip or tongue tip and blade ([ð, l, r, t, n, s, ɾ]). In addition to Catalan, this trend has been found to hold for English ([ʃ] > [ð, d, ɾ], Fowler & Brancazio, 2000), German ([ʃ] > [l, d, n, s, ñ], Hoole, Groerer, & Tillmann, 1990) and Australian languages ([c]>[l], Butcher & Tabain, 2004). Likewise, among vowels, the palatals [i, e] should be more resistant and aggressive than the non-palatals [a, o, u], as also found for [i] vs [a, u] in Taiwan Mandarin (Chen et al., 2015). Moreover, data for Catalan reported elsewhere (Recasens & Espinosa, 2009) suggest that [o, ñ] could be less resistant than [a] perhaps since the formation of a velar or upper pharyngeal constriction for high or mid back rounded vowels constrains the front dorsum to a lesser extent than the formation of a lower pharyngeal constriction for low vowels, mostly so when [a] is especially front.

(b) Regarding the role of consonant manner of articulation, several issues related to coarticulation resistance and aggressiveness deserve to be looked into. In particular, the fricative [s] and trill [ɾ] are expected to be more resistant and aggressive than [ð, l, r, t, n] among dentoalveolars, and the same would hold for the fricative [ʃ] than for [ñ] among palatals. Moreover, we will also look into whether an increase in degree of constraint for [s, r] renders these consonants equally resistant as [ñ, n, ñ] and if trilling and frication have as much of an impact on contextual effects at the back of the vocal tract as they do at the tongue front and dorsum. Regarding dentoalveolars of other manners of articulation, the approximant [ð] should be the least resistant of all consonants, which probably
would not occur if it were a fricative (Fowler & Brancazio, 2000). On the other hand, it may be the case that requirements to lower the tongue sides cause the tongue body to be somewhat more constrained for laterals than for non-laterals. Indeed, transverse compression of the tongue even in clear varieties of /l/ may cause some tongue body backing in addition to some jaw lowering and contribute to an increase in tongue body constraint albeit not as much as for dark /l/ for the production of which the tongue dorsum is subject to considerable lowering and backing (Lindblad & Lunqvist, 2003; Sproat & Fujimura, 1993; Archangeli & Berry, 2010). A different coarticulatory relationship ought to take place between the alveolopalatal lateral [ɾ] and the other palatal consonants [n] and [ʎ]: while the tongue body must remain in a relatively high and front position for the three consonants, the formation of lateral openings for the lateral results in a more anterior constriction location and less dorsopalatal contact, which should result in more vowel coarticulation and thus less coarticulatory resistance (see also Recasens (1999)).

In sum, the following coarticulation resistance hierarchies are expected to hold for the ten consonants and five vowels under analysis in the present study: [ɾ, n, j] > [s, r] > [l, r, t, n] > [ό]; [i,e] > [a, o, u]. Trilling or frication could render [s, r] as resistant as palatal consonants and, among the segmental groups [ɾ, n, j], [l, r, t, n] and [a, o, u], an increase in coarticulatory resistance may occur as a function of frication ([ɾ]) and laterality ([j]) though not ([ɾ]) in the case of consonants and of low vs back rounded vowels. While these predictions are in line with previous coarticulation studies there may be coarticulatory differences among consonants and vowels at different tongue regions which have not been reported earlier.

2. Method

2.1. Recording procedure

Ultrasound recordings were carried out of all possible VCV symmetrical combinations (50) composed of the Catalan consonants [ð, l, r, t, n, s, r, ɾ, n, j] and vowels [i, e, a, o, u]. Ultrasound data for the consonant [k] in the five vowel contexts were also recorded and analyzed for other purposes; thus, as pointed out in Section 2.2.2, data for [ki] and [ku] were used for determining the boundary between the velar and pharyngeal zones on the lingual splines. The VCV sequences were inserted in the Catalan carrier sentence ‘Sap___poc’ “He/she knows___little” where V1 and V2 occur next to [p] and therefore are assumed not to be affected by the other flanking consonants in terms of lingual configuration. VCV sequences always syllabify (V)(CV) in Catalan.

Five native Catalan speakers, i.e., two men (DR, MO) and three women (ES, JU, IM), of 30–60 years of age took part in the recording session. They recorded the sentence list at a comfortable rate six times except for speakers JU (seven tokens) and DR (four tokens). Subjects were asked to utter the two syllables of the VCV sequences of interest with similar degrees of stress. Duration measures taken from spectrographic/waveform displays (see Section 2.2.1) revealed that V1 was slightly longer than V2 for four out of the five speakers, i.e., mean vowel duration across speakers was 72.52 ms for V1 (sd=9.88) and 66.62 ms for V2 (sd=8.70). Overall, each recording session lasted from 45 min to one hour.

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 directly as raw pre-scan converted data using a probe with a 90% field of view and a frequency of 2 MHz in the case of all subjects. The recording sampling rate was 57 frames per second, yielding one image every 17.54 ms, which allowed us to capture one or more tongue contours for all intervocalic consonants since their duration, as determined from the acoustic signal, exceeded 25 ms even for the shortest segment, the tap [ɾ], whose mean duration across speakers was 28.7 ms, and the approximant [ð] (54.2 ms).

Throughout the entire recording session the ultrasound probe was attached to a transducer holder which was positioned under the subject’s chin in an Articulate Instruments Stabilization Headset. Image streams were recorded synchronously with the audio signal sampled at 22,050 Hz. Contours mostly of the back of the alveolar zone and front palate were also recorded by asking speakers to press the tongue against their hard palate or during dry (saliva) swallows (Epstein & Stone, 2005).

Fig. 1. Ultrasound image showing a tongue contour for speaker ES with the subdivision into articulatory zones superimposed. The front of the vocal tract is on the right of the graph.
Fig. 1 plots one of the ultrasound images subject to analysis with the subdivision of the tongue contour into regions, which correspond to the articulatory zones referred to in Section 2.2.2.

2.2. Data analysis

2.2.1. Tongue contour tracking, segmentation and spline manipulation

Tongue contours were tracked automatically at all temporal frames along the entire VCV sequence using the Articulate Assistant Advanced (AAA) software and adjusted manually by the first author whenever necessary. After tracking all the lingual contours, the first author checked them a second time and made additional changes when needed. As noted in the ultrasound literature (Stone, 2005), the mandible and hyoid bones may refract the sound before it reaches the tongue surface thus creating a black region at both edges of the image where the tongue tip and the tongue root are located. This problem was apparent for speaker IM for whom the spline contours could not be traced towards the base of the pharyngeal cavity due to the shadow cast by the hyoid bone (see Section 2.2.2 in this respect). Data points for the tongue contours were exported in an ASCII-file as x-y coordinates with their origin located at the bottom-left corner of the ultrasound image and thus towards the rear of the vocal tract. Acoustic files were also exported in.wav format in order to take segmental duration and spectral measures.

Tongue configuration measurements were performed at the midpoint of V1, C and V2 after identifying V1, C and V2 onsets and offsets on the spectrographic/waveform displays. V1 onset and V2 offset were taken to occur at the onset and offset of vowel-related formant structure, respectively. The acoustic boundary between the vowels and the intervocalic consonant was identified with the beginning and end of several representative events for consonants of different manners of articulation: the frication noise for fricatives; a low intensity formant structure for laterals, nasals and approximants (F1 and F2 spectral discontinuities were also taken into account for segmentation in this case); one short closure for the tap and two or more short closures for the trill. As for the VCV sequences with the stop [t], which is unaspirated in Catalan, V1 offset and C onset were set at the offset of the vowel formants where closure starts, C offset at closure offset immediately before the stop burst and V2 onset at the first noticeable pitch pulse following the stop burst.

Several methodological steps were taken in order to draw the appropriate comparisons at equivalent positions on the tongue surface. Firstly, tongue spline data points were converted from Cartesian coordinates with horizontal coordinate x and vertical coordinate y into polar coordinates with angular coordinate h and radial coordinate r by shifting the origin of the ultrasound image to approximately the center of the ultrasound probe which occurred at X = 86.7 mm and Y = 0 mm in the case of the experimental setup used in the present investigation. This operation is justified by the fact that the tongue surface typically approximates an arc more closely than it approximates a horizontal line (Mielke, 2015). Secondly, smoothing SSANOVA computation was applied using the R package gss to find a best fit curve (Davidson 2006) to the splines of all tokens of each VCV sequence after performing the following edge correction procedure:

(a) Separately for each consonant and vowel we measured the angles delimited by the right and left edges of each available VCV spline and the origin of the ultrasound probe. Then, of all angles measured, we chose the smallest right edge angle and the largest left edge angle. Fig. 2 (left graph) illustrates this procedure for the 20 VCV splines for consonant [l] for speaker DR with angle α being the smallest angle at the right edge of the splines and angle β the largest angle at the left edge.

(b) Of all angles selected in step (a) (i.e., one α angle and one β angle for each consonant and each vowel) we entered in the SSANOVA formula for computing single smoothed splines for each VCV sequence the radian values for the largest of the right edge angles and the smallest of the left edge angles. Thus, in the case of speaker DR the selected angles at the right edge of the splines

![Fig. 2.](image-url)
ranged between 38.07° (the $\alpha$ angle for $V_2$-$[u]$) and 62.2° (the one for $[j]$) while the selected angles at the left edge of the splines ranged between 138.2 degrees (the $\beta$ angle for $[r]$) and 142.12° (the one for $[ɾ]$). Consequently the angles $\alpha$ and $\beta$ entered in the SSANOVA measured 62.2 and 138.2 degrees, respectively. The application of this second step was required since for the SSANOVA smoothing procedure to work properly there needs to be $x$ point values available for at least one token of each VCV sequence.

The SSANOVA smoothed VCV splines consisted of strings of the same number of points separated by 0.01 rad with the associated standard errors corresponding to the radial coordinate variability at each point. Fig. 2 (right graph) plots the smoothed SSANOVA splines for the sequences [ili, ele, ala, olo, ulu] for speaker DR showing the same length and number of points and the $x$ standard errors.

### 2.2.2. Subdivision into articulatory zones

In order to quantify coarticulatory resistance at different articulatory zones, the length of the splines was divided into four portions corresponding to the alveolar, palatal, velar and pharyngeal zones separately for each subject by applying the following criteria.

(a) The boundary between the alveolar and palatal zones was set at the approximate constriction location for the alveolar trill as determined visually by a tongue front inflection point occurring on the $[VrV]$ spline curves displayed in Cartesian coordinates. This articulatory event was selected based on EPG data showing that in Catalan $[r]$ is articulated at rows 3 and 4 of electrodes on the Reading artificial palate and thus at the back alveolar zone (Recasens, 2014b).

(b) The boundary between the palatal and velar zones was taken to occur at closure location for $[k]$ in the sequence $[iki]$ which, also according to EPG data for several Catalan speakers, is found consistently at row 8 of electrodes on the Reading artificial palate and thus at the postpalatal zone just in front of the soft palate. This articulatory event was set at the mean Y maximum across tokens of $[iki]$ as determined on spline curves plotted in Cartesian coordinates.

(c) The length of the velar zone was taken to be 1.26 and 1.51 times that of the palatal zone in the case of the male and female speakers, respectively (Fitch & Giedd, 1999). These ratios correspond to subjects of 19–25 years of age and should not change much later in life. Indeed, Xue & Hao (2003) have shown that, while oral cavity length increases with age, overall vocal tract length and pharyngeal volume do not. The location along the horizontal axis of the Y maximum for $[k]$ was also computed across tokens of the sequence $[uku]$ so as to make sure that the velar closure for a true velar stop was located well inside the velar zone as determined by applying the criterion described above. As expected, this Y maximum value happened to fall about the middle of the velar zone for speakers DR, MO, ES and JU. Applying the female’s ratio to the female speaker IM caused closure location for the velar stop in the sequence $[uku]$ to occur near the velar/pharyngeal boundary, and her velar zone to be too long relative to the palatal zone (for the other four subjects the velar and palatal zones showed a similar length or else the velar zone was slightly longer than the palatal zone). In view of this inconsistent scenario, the male’s ratio was applied to this female subject so as to render her productions of $[k]$ in the context of $[u]$ truly velar and the relative length of the four articulatory zones comparable to those of the other four speakers.

![Fig. 3. Articulatory delimitations superimposed on the SSANOVA lingual splines for $[l]$ in the sequences [ili, ele, ala, olo, ulu] (left) and polygon for the palatal zone shaded in grey (right). PHAR stands for pharyngeal, VEL for velar, PAL for palatal and ALV for alveolar. The front of the vocal tract is on the right of the graph.](image-url)
Fig. 3 (left graph) shows the same smoothed splines displayed in Fig. 2 (right graph) with the pharyngeal, velar, palatal and alveolar articulatory zones superimposed. Table 1 shows the total number of x points on the lingual splines and the number of points at each articulatory zone for each speaker; the smaller number of data points for speaker IM than for the other four subjects was due to difficulties in tracking the tongue surface at the lower pharyngeal zone (see Section 2.2.1). To our knowledge this is the first time that this tongue region division technique has been applied to tongue splines obtained with ultrasound. It should be kept in mind that the subdivision criteria for delimiting the alveolar and palatal zones are based on considerations about constriction location for the consonants of a particular language and are thus language-specific.

### 2.2.3. Coarticulatory resistance evaluation

Coarticulatory resistance and aggressiveness were evaluated for each consonant and vowel. Coarticulatory resistance is inversely dependent on the degree of contextual variability during the target segment. Consonant coarticulatory resistance was computed for each consonant (10 consonants) as a function of all contextual vowels (5 vowels) at the consonant midpoint, and vowel coarticulatory resistance was calculated for each vowel (5) as a function of all contextual consonants (10) at the midpoint of V1 and V2. Coarticulatory aggressiveness is inversely dependent on the degree of contextual variability for each target segment during the flanking segment. Coarticulatory aggressiveness was thus measured for each consonant (10) across all vowels (5) at the midpoint of V1 and V2, and for each vowel (5) across all consonants (10) at consonant midpoint.

Both the degree of coarticulatory resistance and aggressiveness for a given consonant or vowel were determined based on the area of polygons embracing all contextual splines at a given articulatory zone, as delimited by the maximal and minimal Y values at each point on the horizontal axis. The criterion for identifying the articulatory areas is exemplified in Fig. 3 (right graph) showing shaded in grey the polygon encompassing the splines for the five [VIV] sequences with vowels [i, e, a, o, u] at the palatal zone. The area of the polygons was computed using the Gauss determinant algorithm. The smaller the area of the polygon at a given zone, the more coarticulation resistant/aggressive the consonant or the vowel at that zone.

This computation method was found to be more appropriate than the method which estimates coarticulatory resistance for a given consonant or vowel from the area of polygons whose vertices are located at the highest point of the tongue dorsum on the spline curves (Proctor, 2009). Regarding the analysis of coarticulatory resistance for consonants as a function of the contextual vowels, the identification of the highest tongue dorsum point is generally representative of the vowel gesture for the contextual (mid) high vowels but not necessarily so for the contextual low vowels. Thus, the highest tongue dorsum position during the dentoalveolar stop in a sequence like [ata] may occur at the tongue blade rather than at the tongue dorsum and therefore be associated with the tongue gesture for the consonant rather than with that for the vowel. Moreover, this method becomes utterly impractical for evaluating coarticulatory resistance for a target vowel as a function of a considerable number of consonants as we do in Section 3.1.2.

The area values of the polygons for each target consonant and vowel were submitted to a speaker-dependent normalization procedure so as to perform the appropriate comparisons across subjects. For each speaker and articulatory zone, the mean area value across all consonants or vowels was subtracted, respectively, from the area value of each individual consonant or vowel, and the outcome was divided by the standard deviation of that mean. Then, the resulting negative numbers were rendered positive by adding a positive value to all area values.

In order to achieve a better estimate of contextual variability and thus of the degree of coarticulatory resistance and aggressiveness we also computed the coefficients of variation (i.e., the ratio of the standard deviation to the mean) for each target consonant and vowel across contextual conditions at all spline data points along each articulatory zone. It should be emphasized that the two methods for estimating contextual variability used in the present study differ in that, while there are as many coefficients of variation values as spline points, the polygon areas are expressed as a single value encompassing all spline points within a given articulatory zone.

The relationship between coarticulatory resistance and aggressiveness was established based on the hierarchy of segmental units for the two coarticulation types and on results from correlation and regression analyses between polygon area values. As hypothesized in the Introduction section, the measures of coarticulatory resistance and coarticulatory aggressiveness for a given consonant or vowel were expected to be highly positively correlated. Thus, for example, palatal consonants were expected to be more resistant than several dentoalveolar consonants to the contextual vowel effects and at the same time to exert more prominent coarticulatory effects during the surrounding vowels.

<table>
<thead>
<tr>
<th>Speaker</th>
<th>Overall splines</th>
<th>PHAR</th>
<th>VEL</th>
<th>PAL</th>
<th>ALV</th>
</tr>
</thead>
<tbody>
<tr>
<td>DR</td>
<td>133</td>
<td>54</td>
<td>33</td>
<td>29</td>
<td>17</td>
</tr>
<tr>
<td>MO</td>
<td>125</td>
<td>56</td>
<td>28</td>
<td>26</td>
<td>15</td>
</tr>
<tr>
<td>ES</td>
<td>131</td>
<td>52</td>
<td>33</td>
<td>29</td>
<td>17</td>
</tr>
<tr>
<td>JU</td>
<td>124</td>
<td>61</td>
<td>29</td>
<td>20</td>
<td>24</td>
</tr>
<tr>
<td>IM</td>
<td>101</td>
<td>28</td>
<td>30</td>
<td>25</td>
<td>19</td>
</tr>
</tbody>
</table>
2.2.4. Statistics

Linear mixed model (LMM) tests were performed on the normalized area values at each articulatory zone with ‘subject’ \(k\) in the statistics formulas shown below as a random factor using the ‘mixlm’ package of R version 3.1.2 (R Developmental Core team 2014). The LMM design was applied in line with the fact that, as verified for each test, the residual standard errors for the current dataset turned out to be fairly normally distributed. Significant differences in coarticulatory resistance and aggressiveness were obtained applying the models specified in the formulas A, B, C, D below, where \(\alpha, \beta\) and \(\delta\) stand for the fixed factors, \(\gamma, \zeta, \lambda\) and \(\eta\) for the factor interactions, \(B\) for the random factor and \(e\) for the random error. The statistical models include the fixed factors consonant \((i)\) with levels \([\delta, l, r, t, n, s, r, \lambda, \j, \] and zone \((j)\) with levels alveolar, palatal, velar and pharyngeal in the case of coarticulatory resistance for consonants (formula A), and vowel \((n)\) with levels \([i, e, a, o, u]\), zone \((j)\) with levels alveolar, palatal, velar and pharyngeal, and vowel position \((r)\) with levels V1 and V2, for the test performed on coarticulatory resistance for the vowel data (formula B). As for coarticulatory aggressiveness, the fixed factors were ‘consonant’, ‘zone’ and ‘vowel position’ for the consonantal data (formula C), and ‘vowel’ and ‘zone’ for the vowel data (formula D). The significance level was set at \(p<0.05\), and Tukey post-hoc tests were executed whenever main effects and two- or three-factor interactions achieved significance.

\[
\text{(A)} y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + B_k + e_{ijk}
\]

\[
\text{(B)} y_{njk} = \mu + \alpha_n + \beta_j + \delta_r + \gamma_{nj} + \lambda_r + \eta_{nr} + B_k + e_{njrk}
\]

\[
\text{(C)} y_{ijk} = \mu + \alpha_i + \beta_j + \delta_r + \gamma_{ij} + \lambda_{jr} + \eta_{ijr} + B_k + e_{ijrk}
\]

\[
\text{(D)} y_{njk} = \mu + \alpha_n + \beta_j + \gamma_{nj} + B_k + e_{njrk}
\]

The statistical results for the normalized dataset will be interpreted with reference to the ‘consonant’ and ‘vowel’ main effects and to the consonant- and vowel-dependent differences across zones and within each articulatory zone, which is the main topic of our investigation. Statistical results for the ‘zone’ and ‘vowel position’ main effects will not be given since the normalization procedure happened to level out the differences in area size among the polygons for the four articulatory zones and the two vowel locations (V1, V2). In order to obtain information about these two measures, we ran the same statistical analyses described above on the unnormalized data values and will present the corresponding main effects and interactions in the Results section.

Fig. 4. Results from Tukey post-hoc tests for the ‘consonant’ main effect (top) and the ‘consonant’ x ‘zone’ interaction (bottom) performed on the normalized area values for consonants measured at consonant midpoint. PHAR stands for pharyngeal, VEL for velar, PAL for palatal and ALV for alveolar.
3. Results

3.1. Coarticulatory resistance

3.1.1. Consonants

The statistical test performed on the normalized area size values for the consonant resistance condition yielded a main effect of ‘consonant’ (F(9, 156) = 79.45, p < 0.001), ‘zone’ (F(3, 156) = 3.93, p < 0.05) and a significant ‘consonant’ x ‘zone’ interaction (F(27, 156) = 3.05, p < 0.001). Consonant-dependent differences in coarticulatory resistance correspond to the significant effects resulting from the post-hoc tests for the ‘consonant’ main effect and represented in grey for the relevant two-consonant pair combinations in Fig. 4, top graph. The sign of the significant differences may be found by inspecting the normalized area sizes for the different consonants in Fig. 5 irrespective of articulatory zone. Overall, contextual variability decreases in the progression [ð] (most variable, least resistant) > [l, r, t, n] > [s, r] > [ʌ, i, n] (least variable, most resistant). Thus, at the two extremes we find [ð] and the three palatals which differ significantly from all other consonants and have the largest and smallest polygon areas, respectively. Regarding the consonants showing intermediate area values, [s, r] have significantly smaller polygons than [l, r, t, n] as a general rule (except for [s] vs [n]) and consonants within each of the two groups do not differ among themselves.

Consonant-dependent differences in area size appearing in Fig. 5 achieved significance at all four articulatory zones in ways shown in the four bottom graphs of Fig. 4 and summarized in (a) and (b) below.

(a) As for the palatal, velar and pharyngeal zones where the tongue body is located, the palatal consonants [ʌ, ɾ, ʃ] were significantly less variable and thus more resistant than [ð, l, r, t, n]. Moreover, among [ð, l, r, t, n] the approximant [ð] is the least resistant consonant since it may differ significantly from [n, s, r] while [l, r, t, n], which have smaller area values than the dental approximant, are not set in contrast at any of the three zones.

The relationship between [s], and to some extent [ɾ], and the remaining consonants differs depending on the zone taken into consideration. More specifically, the area values of these two consonants become more similar to those for the other dentoalveolars and increasingly different from those for the three palatals as the back vocal tract is approached, as indicated by several significant differences which were absent at the palatal zone and present either at the velar and pharyngeal zones (between [s] and [ʌ, ɾ, ʃ]) or just at the pharynx (between [ɾ] and [n]). Overall, [s] and [ɾ] occupy an intermediate position between those consonants which are most resistant and least resistant, and [s] and less so [ɾ] become more sensitive to vowel effects at the rear of the vocal tract than at the palatal zone.

(b) The statistical results reported in Fig. 4 (bottom right graph) and the mean area values presented in the rightmost panel of Fig. 5 reveal some consonant-dependent differences in polygon area size at the alveolar zone, i.e., significantly larger polygons and thus more coarticulation for [ð, r] than for [s, ɾ, ʃ] (also for [l] than for [ʃ]). Maximal vowel coarticulation for [ð] and [ɾ] may be due to the manner of articulation characteristics involved in the production of the two consonants: they are the shortest of all consonants subject to investigation (Section 2.1) and articulated with the tongue tip, which is more prone to adapt to phonetic context due to the manner of articulation characteristics involved in the production of the two consonants: they are the shortest of all consonants subject to investigation (Section 2.1) and articulated with the tongue tip, which is more prone to adapt to phonetic context.

This coarticulatory resistance scenario becomes apparent when the lingual configurations for the ten consonants in all five vowel contexts are taken into consideration. Lingual splines for speaker DR reveal that there is lesser vowel coarticulation at the tongue blade and dorsum and at the tongue back for the three palatals (Fig. 6b) than for the seven dentoalveolars (Fig. 6a) as a general rule. Vowel effects on the dentoalveolar consonants are exerted by [i, e] vs [a, o, u] both in tongue dorsum raising at the tongue blade and dorsum and at the tongue back for the three palatal consonants (Fig. 6b) than for the seven dentoalveolars (Fig. 6a) as a general rule (except for [s] vs [n]) and consonants within each of the two groups do not differ among themselves.

In sum, coarticulatory resistance for consonants varies in the progression [ʌ, ɾ, ʃ] (most resistant) > [s, r] > [l, r, t, n] > [ð] (least resistant) at the palatal, velar and pharyngeal articulatory zones, with the partial exception of [s] and to a lesser extent [ɾ] which may be as variable as [l, r, t, n] at the back of the vocal tract. Regarding the alveolar zone, the apical tap and the dental approximant are most variable and the three laminodorsal consonants [ʌ, ɾ, ʃ] and [s] the least variable.
As pointed out in the Method section, the ANOVA performed on the unnormalized area size values provides information about differences in size among articulatory zones across consonants and for each individual consonant. Statistical results in this case yielded a main effect of ‘consonant’ ($F(9,156) = 23.52, p < 0.001$) and ‘zone’ ($F(3, 156) = 254.19, p < 0.001$) and a ‘consonant’ x ‘zone’ interaction ($F(27, 156) = 2.26, p < 0.001$). In parallel to the area sizes for the consonant [l] displayed in Fig. 3, all zone-dependent differences achieved significance and proceeded in the progression pharyngeal>velar>palatal>alveolar. Moreover, area size differences among consonants were significant in some cases: differences were significantly larger for non-palatals than for palatals at the pharyngeal zone and for [ð] than for palatals at the palatal and velar zones.

As referred to in Section 2.2.3, additional information about consonant-dependent differences in contextual variability was derived from coefficient of variation values computed across contextual vowels at successive spline data points. Coefficient of variation curves at the palatal and pharyngeal zones are shown in Fig. 7a for the relatively unconstrained dentoalveolar consonants [ð, l, r, t, n] and for the highly constrained palatals [k, n, j], and in Fig. 7b for the more constrained alveolars [s, t]. Data are shown only at these two zones since they exhibit the most clear coarticulatory patterns and also because, at least for a subset of phonetic segments, changes in tongue dorsum raising at the palatal zone are correlated with variations in tongue body fronting at the back of the vocal tract. Coefficient of variation values at the consonant midpoint plotted in Fig. 7a (five left graphs) reveal that, analogously to the area...
values reported in Section 3.1.1(a), the palatals $\emptyset$, $\emptyset$, $\emptyset$] are less variable and thus more resistant than the dentoalveolars $\emptyset$, $\emptyset$, $\emptyset$, $\emptyset$, $\emptyset$] both at the palatal zone and at the pharynx in the case of all subjects. Generally speaking, the degree of vowel-dependent variability is higher at the pharynx than at the palate, while the contrast between the two consonant groups is better defined at the latter zone than at the former. Moreover, there is a general trend for contextual variability to increase as we proceed from the front to the back location within the pharyngeal zone, which means that the vowel-dependent differences in tongue position are larger towards the base of the pharyngeal region than at its upper edge. The coefficient of variation curves at the midpoint of $[s]$ and $[r]$ (Fig. 7b, five left

Fig. 7. a. Coefficient of variation values at all spline points along the pharyngeal (PHAR) and palatal (PAL) zones for consonants $[\emptyset, \emptyset, \emptyset, \emptyset]$ and $[\emptyset, \emptyset, \emptyset]$ across vowel contexts according to speakers DR, ES, JU, IM and MO. Data correspond to the midpoint of the consonantal period (left) and of V2 (right). The front of the vocal tract is on the right of the graph. b. Coefficient of variation values at all spline points along the pharyngeal (PHAR) and palatal (PAL) zones for consonants $[s, r]$ across vowel contexts according to speakers DR, ES, JU, IM and MO. Data correspond to the midpoint of the consonantal period (left) and of V2 (right). The front of the vocal tract is on the right of the graph.

Fig. 8. Normalized area values for vowels across consonant contexts and speakers at the four articulatory zones measured at V1 and V2 midpoint. This measure is inversely related to coarticulatory resistance for vowels. Vowels in the caption are arranged following the same order as in the graph bars.

values reported in Section 3.1.1(a), the palatals $[\emptyset, \emptyset, \emptyset]$ are less variable and thus more resistant than the dentoalveolars $[\emptyset, \emptyset, \emptyset, \emptyset]$ both at the palatal zone and at the pharynx in the case of all subjects. Generally speaking, the degree of vowel-dependent variability is higher at the pharynx than at the palate, while the contrast between the two consonant groups is better defined at the latter zone than at the former. Moreover, there is a general trend for contextual variability to increase as we proceed from the front to the back location within the pharyngeal zone, which means that the vowel-dependent differences in tongue position are larger towards the base of the pharyngeal region than at its upper edge. The coefficient of variation curves at the midpoint of $[s]$ and $[r]$ (Fig. 7b, five left
3.1.2. Vowels

The statistical test for the vowel area sizes yielded a main effect of ‘vowel’ (F(4, 156) = 80.59, p < 0.001) and no significant ‘vowel’ x ‘zone’ or ‘vowel position’ x ‘zone’ interactions. As shown in Fig. 8 and according to results from the post-hoc tests, vowels turned out to differ significantly in degree of consonant-dependent variability at their midpoint for [u] (most variable, least resistant) > [o] > [a] > [i, e] (least variable, most resistant), and these differences held constant at V1 and V2 and at all four articulatory zones.

Vowel-dependent differences in tongue body coarticulatory resistance may be inferred from the lingual configurations for [i, e, a, o, u] in all consonant context conditions sampled at the midpoint of V1 and V2. As shown by the ten lingual splines for each of the five vowels at V2 midpoint in Fig. 9 there is less consonantal coarticulation at the tongue dorsum and tongue back for the palatal vowels [i, e] than for the low vowel [a] and the back rounded vowels [o, u]. The figure also shows that consonant effects on [a, o, u] occur mostly in tongue dorsum raising and tongue postdorsum fronting as a function of the palatals [k, p, j] vs the non-palatals [ð, l, r, t, n, s, r].

Statistical results for the unnormalized area size values yielded a main effect of ‘vowel’ (F(4, 156) = 31.05, p < 0.001), ‘zone’ (F(3, 156) = 156.36, p < 0.001) and ‘vowel position’ (F(1, 156) = 160.42, p < 0.001), and the significant interactions ‘vowel’ x ‘zone’ (F(12, 156) = 2.95, p < 0.001), ‘vowel’ x ‘vowel position’ (F(4, 156) = 3.47, p < 0.05) and ‘zone’ x ‘vowel position’ (F(3, 156) = 31.13, p < 0.001). The triple interaction ‘vowel’ x ‘zone’ x ‘vowel position’ did not achieve significance. Differences in area size among zones were analogous to the ones reported for consonants (Section 3.1.1), i.e., pharyngeal>palatal, velar>alveolar, though, while all vowels exhibited greater polygon areas at the pharynx than at the remaining zones, this zone-dependent effect was more obvious for back rounded vowels than for front vowels. Moreover, consonant-dependent variability was greater at V2 than at V1 due, among other possible factors, to the fact that the tongue configuration for low and back rounded vowels was especially sensitive to the coarticulatory action of preceding palatal consonants (see above in this same section).

Coefficient of variation curves at the palatal and pharyngeal zones for the five vowels plotted in Fig. 10 (five left graphs) are largely consistent with the results from the LMM analysis. As expected, contextual variability at the palatal zone is less for [i, e] than for [a, o, u] for all speakers and for [a] than for [o, u] for four speakers. A similar pattern holds at the pharynx for speakers IM and MO, but not clearly so for the other three speakers, who may exhibit considerable contextual variability not only for low and back vowels but for front vowels as well (see, for example, the coefficient of variation curves for [i] for subjects ES and JU). This latter outcome is not clearly corroborated by the polygon area values though some increase in contextual variability for [i] at the pharyngeal and velar zones vis-à-vis the other two articulatory zones is apparent in the left graph of Fig. 8. A comparison between the Figs. 10 and 7a and b also shows that contextual variability for vowels does not increase as we approach the base of the pharynx or else that whenever there is such an increase it happens to be less obvious than for consonants.

3.2. Coarticulatory aggressiveness

3.2.1. Consonants

Statistical results for consonants at the vowels midpoint yielded a main effect of ‘consonant’ (F(9, 316) = 50.98, p < 0.001) and significant ‘consonant’ x ‘zone’ and ‘consonant’ x ‘vowel position’ interactions (F(27, 316) = 3.13, p < 0.001; F(9, 316) = 3.63, p < 0.001). Overall consonant-dependent differences were similar to those reported at the midpoint of the consonantal period, though
Fig. 10. Coefficient of variation values at all spline points along the pharyngeal (PHAR) and palatal (PAL) zones for vowels [i, e, a, o, u] across consonant contexts according to speakers DR, ES, JU, IM and MO. Data correspond to the midpoints of V2 (left) and the consonant (right). The front of the vocal tract is on the right of the graph.
the number of significant effects was smaller (compare the statistical differences in the top graphs of Figs. 11 and 4, on the one hand, and also the mean normalized area values in Figs. 12 and 5, on the other).

Regarding the ‘consonant’ x ‘zone’ interaction, consonant-dependent differences turned out to be highly significant at each of the four articulatory zones. Moreover, in contrast to the statistical results reported in the bottom graphs of Fig. 4, the area values for [r, ʎ] reported at the palatal, velar and/or pharyngeal zones at the bottom of Fig. 11 (see also Fig. 12) show more affinity with those for [n, j]. This means that the two consonants of interest exhibit a more variable tongue body configuration when measured at the vowel midpoint than during the consonant. As Figs. 11 and 12 also show, consonant-dependent differences at the alveolar zone hold essentially between the apicals [ð, r] and [s, ɲ, j] (also between [n] and [s, j]) and thus conform to a large extent to the differences in coarticulatory sensitivity occurring at the consonant midpoint.

The ‘consonant’ x ‘vowel position’ interaction, on the other hand, turned out to be related mostly to the presence of more consonant-dependent variability at V1 than at V2 in VCV sequences with [ʎ] and at V2 than at V1 in those with [l] (see Fig. 12). These position-dependent differences appear to be associated with more gestural anticipation than carryover for the alveolar lateral and the reverse for the alveolo-palatal lateral; thus as the anticipatory C-to-V effects increase (for VCV sequences with [ll]) vowels become less
distinct at V1 midpoint while an increase in the strength of the carryover C-to-V effects (for VCV sequences with [ʎ]) cause the vowel-dependent differences at V2 midpoint to become smaller. Consonant-dependent differences in area size at the two flanking vowels were found to be highly correlated with those occurring at the consonant midpoint, which is in support of the notion that the degree to which consonants affect adjacent vowels is determined by how resistant they are. Moreover, as shown in Table 2 (top), Pearson correlation coefficient values for the area sizes measured at the pharyngeal, velar and palatal zones were higher in the case of the V2 data ($r$ values range between 0.86 and 0.89) than in that of the V1 data ($r = 0.67–0.81$), while lower correlation values were obtained for the area sizes at the alveolar zone. Regression analyses yielded analogous zone- and position-dependent results: the $r^2$ values at the pharynx, velar zone and hard palate were higher at V2 (between 0.74 and 0.79) than at V1 (between 0.44 and 0.66), and lower at the alveolar zone than at the other three zones.

Results for the unnormalized area values yielded a main effect of ‘consonant’ ($F(9,316) = 10.28$, $p < 0.001$), ‘zone’ ($F(3, 316) = 543.32$, $p < 0.001$) and a significant ‘consonant’×‘position’ interaction ($F(9, 316) = 2.52$, $p < 0.05$). Polygon areas differed in size among articulatory zones in the progression pharyngeal>velar, palatal>alveolar, and were larger at V1 midpoint than at V2 midpoint presumably since the former vowel was less sensitive to differences in C-to-V coarticulation as a function of palatal vs non palatal consonants than the latter (see also Section 3.1.2). Moreover, a comparison between the unnormalized values for consonants at the consonant and vowel periods reveals a trend for the former to be somewhat less variable than the latter (at least for the most constrained consonants), which supports the notion that variability for consonants increases as we depart from the consonantal period.

In sum, while consonant-dependent differences in coarticulatory resistance and coarticulatory aggressiveness conform to a similar hierarchy, the latter are somewhat more loosely defined than the former in that consonant-dependent differences are smaller and constrained consonants such as [r, ʎ] become more variable.

A comparison of the right and left graphs in Figs. 7a and b reveals that the coefficient of variation curves at the palatal and pharyngeal zones for consonants at V2 midpoint (coarticulatory aggressiveness) parallel those at the consonant midpoint (coarticulatory resistance). As pointed out for the area sizes above, the coefficients of variation at the midpoint of the vowel are slightly higher than those obtained at the midpoints of the consonant.

### Table 2

<table>
<thead>
<tr>
<th></th>
<th>PHAR</th>
<th>VEL</th>
<th>PAL</th>
<th>ALV</th>
</tr>
</thead>
<tbody>
<tr>
<td>$r$ V1</td>
<td>0.67</td>
<td>0.67</td>
<td>0.81</td>
<td>0.52</td>
</tr>
<tr>
<td>$r$ V2</td>
<td>0.89</td>
<td>0.86</td>
<td>0.89</td>
<td>0.67</td>
</tr>
<tr>
<td>$r^2$ V1</td>
<td>0.44</td>
<td>0.45</td>
<td>0.66</td>
<td>0.27</td>
</tr>
<tr>
<td>$r^2$ V2</td>
<td>0.79</td>
<td>0.74</td>
<td>0.79</td>
<td>0.45</td>
</tr>
</tbody>
</table>

#### 3.2.2. Vowels

Statistical results for vowels at consonant midpoint yielded a main effect of ‘vowel’ ($F(4,76) = 31.11$, $p < 0.001$) and a significant ‘vowel’×‘zone’ interaction ($F(12,76) = 4.14$, $p < 0.001$). As shown in Fig. 13, the ‘vowel’ main effect was related to differences in area size for [u, o, a]>[i, e]. Therefore, the vowel-dependent hierarchy was similar to the one seen in the coarticulatory resistance scenario (see Fig. 8) except for the fact that low and back rounded vowels were non-contrastive. The ‘vowel’×‘zone’ interaction was associated with the presence of those same vowel-dependent differences in area size at the alveolar and palatal zones, with effects

![Fig. 13. Normalized area values for vowels across consonant contexts and speakers at the pharyngeal (PHAR), velar (VEL), palatal (PAL) and alveolar (ALV) zones measured at consonant midpoint. This measure is indicative of coarticulatory aggressiveness for vowels.](image)
from just front vs back rounded vowels at the pharynx, and with no vowel effects at the velar zone.

In parallel to the coarticulatory resistance and aggressiveness scenario for consonants, Table 2 (bottom) reveals the existence of a high positive correlation and high regression values between the vowel-dependent differences in area size at the midpoint of V2 and less so of V1 and those occurring at the midpoint of the consonant when the palatal zone and the pharynx are taken into consideration. Thus, r always exceeds 0.90 for the V2 data and ranges between 0.60 and 0.85 for V1, while $R^2$ amounts to 0.84 and 0.89 in the case of the V2 data and to less than 0.7 for V1. Correlation and regression values were lower at the velar and alveolar zones ($r=0.5–0.7, R^2=0.3–0.5$).

The statistical tests for the unnormalized dataset yielded a main effect of vowel ($F(4,76)=12.95, p<0.001$) and ‘zone’ ($F(3,76)=104.66, p<0.001$) and a barely significant ‘vowel’ x ‘zone’ interaction ($F(12,76)=1.89, p<0.05$). The zone-dependent effect was related to differences in area size decreasing in the progression pharyngeal > velar, palatal > alveolar and to larger consonant-dependent differences for palatal vowels than for back labial vowels at the pharynx (see also Section 3.2.1). Variability for vowels was greater at the consonant midpoint than at the vowels midpoint, which provides support for the notion that vowels should be affected by consonants to a larger extent away from the vowel period than during the vowel.

A comparison of the right and left graphs of Fig. 10 reveals that the coefficient of variation curves for the five vowels at the palatal and pharyngeal zones conform to a similar hierarchy whether computed at the intervocalic consonant (coarticulatory aggressiveness) or at the vowel midpoint (coarticulatory resistance).

4. Discussion

Data on articulatory area size and coefficient of variation values reported in this study reveal differences in coarticulatory resistance for front lingual consonants and vowels varying in the progression [ʃ, n, ɿ] (most resistant) > [r] > [s] > [l, t, n] > [ʊ] (least resistant) and [ɪ, e] > [a] > [o] > [ʊ]. Regarding these coarticulatory resistance hierarchies it needs to be emphasized that [ɪ] is clear rather than dark among the Catalan speakers of the present study and that the phonetic symbol [ʊ] corresponds to an approximant realization of /d/. These consonant and vowel hierarchies are largely in agreement with data on tongue dorsum coarticulatory resistance for Catalan consonants and vowels reported in earlier EPG and EMA studies. They also lend support to the degree of articulatory constraint (DAC) model of coarticulation in that the extent to which a portion of the tongue body is more or less resistant to coarticulation depends both on whether this articulatory structure is or is not involved in the formation of a closure or constriction and on the severity of the manner of articulation requirements (Recasens et al., 1997). In particular, the tongue body is more constrained for palatal segments than for consonants articulated further forward in the mouth and for vowels showing a more retracted constriction location. Moreover, among dentoalveolar consonants, lingual fricatives and trills happen to be more constrained than consonants of other manners of articulation.

Another relevant finding of the present investigation is that the consonant and vowel hierarchies in question hold at the palatal, velar and pharyngeal zones where the tongue body is located, and not just at about the palatal zone as reported in previous studies using other methods for tracking front lingual activity and lingual contact (Recasens et al., 1997; Recasens & Espinosa, 2009). Thus, vowel-to-consonant coarticulation was less for palatal consonants and palatal vowels (also for the trill [ɾ]) than for the less constrained dentoalveolar consonants [ð, l, r, t, n] at the three articulatory zones in question. In so far as the acoustic output depends on the overall vocal tract cavity configuration, this finding may account for why spectral coarticulation data reported elsewhere exhibit analogous differences in coarticulatory resistance to those found in the present investigation. Data from the present study also reveal that consonant-dependent differences at the alveolar zone are associated with apical (more variable) vs laminal/laminopredorsal (less variable) articulations; in particular, coarticulatory resistance was less present for [ʊ, ŋ] than for the remaining consonants.

More specific findings about coarticulation resistance reported in the present study labeled (a) through (f) below deserve special attention.

(a) While tongue dorsum raising and fronting induced by contraction of the genioglossus muscle constrain the entire tongue body for palatal consonants and palatal vowels, this appears to be the case for the former phonetic segments rather than for the latter. Thus, in contrast with [ɿ, n, ɿ, ɿ] may be more variable at the tongue postdorsum and root than at the front dorsum for some speakers (see also Kiritani, Itoh, Hirose, and Sawashima (1977) regarding Japanese vowels). Less contextual sensitivity at the tongue front than at the tongue root for palatal vowels could be related to the role of the anterior genioglossus whenever the posterior genioglossus is activated, and may account for differences in pharyngeal volume for high vowels specified for the ATR (advanced tongue root) contrast in Akan and other languages of West and East Africa (Tiede, 1996). Albeit less clearly, differences in tongue dorsum height for tense vs lax vowel in English are also correlated with differences in tongue back position and pharyngeal cavity size.

It appears that the tongue positioning demands imposed by speakers during the production of palatal consonants are stricter than those associated with manner of articulation characteristics such as frication and trilling for the dentoalveolar consonants [s] and [ɾ]. In fact, differences in coarticulatory resistance among palatal consonants of different manners of articulations turned out to be rather small which is also consistent with the high degree of tongue body constraint involved in the implementation of the tongue dorsum fronting and raising gesture. In any case, in agreement with linguopalatal contact data on V-to-C coarticulation reported for Catalan in the literature (Recasens, 2014b), those differences occurred between [ɿ] and [n, ɿ] when data for coarticulatory aggressiveness were taken into consideration (i.e., [ɿ] could be less aggressive than [n, ɿ]). A possible reason for the difference in coarticulatory resistance and aggressiveness between [ɿ] and [n] is that the lateral is more anterior than the nasal and produced with less dorsopalatal contact, presumably since one or both sides of the
mouth must remain open for the passage of airflow. A higher degree of coarticulatory resistance for [l] than for [ð, ð] was apparent at the palatal and velar zones (see Fig. 5) but did not achieve significance.

(b) The highly constrained fricative consonant [s] turned out to be more sensitive to vowel effects at the velar and pharyngeal zones than at constriction location (see also Zharkova et al. (2012)). This finding is in line with the fact that the generation of turbulent noise for lingual fricatives depends primarily on the area of the tongue constriction and on the pressure difference between the cavities located in front of and behind the constriction in addition to the passage of sufficient airflow through the glottis (Ohala & Soile, 2010).

These conditioning factors block the vocico effects at the alveolar and palatal zones to a large extent while leaving some room for contextual adaptation at the back of the vocal tract. Moreover, the alveolar trill appears to be more constrained than the alveolar fricative at most articulatory locations which is in accordance with the articulatory and aerodynamic factors involved in the performance of trills, i.e., appropriate tongue tip tension and velocity and sufficient back cavity pressure and airflow, which, if lacking, may result in the absence of tongue-tip vibration (McGowan, 1992).

Little vowel coarticulation for [r] at the velar zone and upper pharynx could result from the active formation of a secondary constriction (Proctor, 2009) or, more likely, from the fact that the tongue body needs to be lowered and retracted in order to perform the trill successfully. In our opinion tongue body backing associated with the manner requirements for the alveolar trill differs from active tongue body retraction associated with the formation of consonants articulated at the back of the vocal tract such as Arabic pharyngeals or pharyngealized apico-dentoalveolars. Thus, in contrast with these Arabic consonants, there is no good reason to assume that a secondary pharyngeal constriction ought to take place during the production of the alveolar trill in Catalan or Spanish since, in combination with some alveolar closure backing and predorsum lowering, the final goal of this tongue back retraction movement is to assist the performance of the tongue tip trilling motion. The fact that rhotics may cause adjacent dental and front alveolar consonants to become postalveolar or retroflex (Bhat, 1973) does not imply that trills are implemented through two lingual gestures but may result naturally from the manner of articulation requirements on tongue configuration.

(c) Coarticulatory resistance was less for [l] than for [s, r] which supports the notion that laterality (i.e., the requirements to lower the tongue sides) contributes less than friction and trilling to an increase in tongue body constraint. The fact that the alveolar lateral was clear rather than dark may account partly for this finding: it is known that an increase in darkness degree causes the tongue body to lower and retract and the alveolar lateral to become more resistant to coarticulation. In any case, we expected [l] to allow less vowel coarticulation than non-lateral alveolars such as [n] and perhaps the fast tap realization [ɾ] since, in combination with some alveolar closure backing and predorsum lowering, the final goal of this tongue back retraction movement is to allow the performance of the tongue tip trilling motion. The fact that fricatives may cause adjacent dental and front alveolar consonants to become postalveolar or retroflex (Bhat, 1973) does not imply that trills are implemented through two lingual gestures but may result naturally from the manner of articulation requirements on tongue configuration.

(d) Among the lesser constrained dentoalveolars, [ð] turned out to be less constrained than [l, r, t, n], presumably since it is produced with a wider and looser lingual constriction. Acoustic data also reveal that the approximant cognates [ð], [ð] and [y] are highly sensitive to vowel coarticulation in languages like Spanish and Catalan (Recasens, 2014b). The glide [j], which may also be treated as an approximant, is exceptional in this respect due to the tongue body positioning demands for palatal sounds referred to above.

(e) As for low and back rounded vowels, differences in coarticulatory resistance for [a] > [o] > [u] suggest that the formation of a lower pharyngeal constriction constrains the tongue dorsum to a larger extent than the formation of a velar or upper pharyngeal constriction. This scenario could result from differences in muscle activation between the two sets of vowels (Buchaijall, Perrier & Payan, 2009, Baer, Alfonso & Honda, 1988, Kakita, Fujimura & Honda, 1985, Wood, 1975). High and mid back rounded vowels involve mainly the activation of the styloglossus muscle, which pulls the tongue body upwards and backwards and, in addition, the posterior genioglossus for [u] and the pharyngeal constrictors for [o]. At least for [u] the entire genioglossus muscle could assist the tongue back raising by relaxation, which could account for the high degree of contextual sensitivity towards the tongue front when the vowel occurs next to palatal and dentoalveolar consonants (see Takano & Honda (2007) for Japanese). Low vowels, on the other hand, are produced by activating not only the hyoglossus which pulls the tongue backwards and downwards and the pharyngeal constrictors, but also the anterior genioglossus which acts to keep the tongue dorsum surface flat and in a relative front position more so when the low vowel is front ([a], also [æ]) than when it is back [o]. Whether because of the latter muscular action or for some other reason, contextual effects on the dorsum of the tongue for [a] are, albeit considerable, somewhat less prominent than those for [u] in Catalan (see also Recasens (1999)). It deserves to be investigated whether this vowel-dependent difference in coarticulatory sensitivity depends on the articulatory characteristics of [a] and may thus vary from language to language.

Another major finding of the present investigation is that coarticulatory resistance and coarticulatory aggressiveness turned out to be positively correlated for both consonants and vowels and that this relationship held at the body of the tongue rather than at the tongue front. Indeed, correlation values between the two coarticulatory dimensions were about 0.7–0.9 at the palatal, velar and pharyngeal constrained zones for consonants, and at the palatal and pharyngeal zones for vowels. This finding is in line with previous remarks in the literature (Fowler & Saltzman, 1993) and supports the general predictions of the DAC model of coarticulation. It implies that the tongue body activity spans over a greater temporal period during the flanking segments for consonants and vowels which are more constrained than for those which are less constrained. Thus, for example, higher levels of tongue body activation at the midpoint of the flanking vowels in VCV sequences with palatal vs dentoalveolar consonants cause the lingual configuration for different vowels to approach that for the intervocalic consonant to a large extent and thus to become less contrastive.
Also regarding the coartulatory aggressiveness issue, V2 was found to be more variable and thus more consonant-dependent than V1 due to the special prominence of the carryover effects exerted by palatal vs non-palatal consonants on low and back rounded vowels. Consequently, correlation values between coartulatory resistance and coartulatory aggressiveness were higher when the V2 data were taken into consideration.

The conclusions of the present investigation are based on a relatively small number of speakers of a single language, which calls for their confirmation in future production studies with more subjects. In spite of this limitation our results are generally consistent with studies reporting differences in coartulatory resistance and aggressiveness among consonants and vowels using other recording and analysis techniques and speakers from other languages (see Section 1).

The present findings have implications for theories of speech production and sound change. Assimilatory processes are predicted to occur mostly in segmental sequences with adjacent unconstrained and constrained segments since the two segment types are most prone to coarticulate with each other, as for example in the case of the sequences [na] (which should exhibit considerable vowel raising and fronting) and [ni] (which is expected to undergo consonant palatalization). Assimilatory effects may also take place whenever a highly constrained consonant is preceded and/or followed by a highly constrained vowel and the two are antagonistic (e.g., VC sequences composed of [r] or dark /l/ and [I]) given that consonants are subject to stricter production requirements than vowels. Moreover, the phonetic outcome of the assimilatory processes in question ought to be related to the lingual region which exhibits a higher degree of constraint and therefore undergoes minimal coarticulation. Thus, for example, [r] may cause vowels to back if front and to lower if high, while [n] participates mostly in [a] raising and fronting and in [o, u] fronting.

Acknowledgments

This research was supported by project FFII2013-40579-P of the Spanish Ministry of Economy and Competitiveness, by the ICREA Academia program, and by project 2014SGR61 of the Catalan Government. We would like to thank three reviewers for comments on a previous version of the manuscript, Alan Wrench for assistance with the ultrasound data recording and analysis procedures, and the audience at the Ultrafest VII conference (Hong Kong, December 8-10, 2015) for relevant questions posed during our presentation.

References

Lindblad, P., & Lundqvist, S. (2003). [t] tends to be velarised, apical as opposed to laminal, and produced with a low jaw, and these features are connected. In M.J. Solé, D. Recasens, & J. Romero (Eds.), Proceedings of the 15th international congress of the phonetic sciences (pp. 1899–1902), Barcelona: Causal.


