Acoustic Characteristics of Velar Stops and Velar Softening in German, Polish and Catalan

Marzena Zygis[1], Daniel Recasens[2] & Aina Espinosa[3]

[1] Zentrum für Allgemeine Sprachwissenschaft, Berlin
[2,3] Universitat Autònoma de Barcelona & Institut d'Estudis Catalans
E-mail: [1]marzena@zas.gwz-berlin.de, [2]daniel.recasens@uab.es

Abstract

Two hypotheses have been proposed in order to account for velar softening, i.e., a process through which /k/ changes to an affricate. Whereas one hypothesis states that for the process to apply the velar stop has to be realized as an (alveolo)palatal stop (articulation-based hypothesis), the other claims that velar softening is triggered by acoustic similarity between the input and output segments (acoustic equivalence hypothesis). The present paper investigates the acoustic equivalence hypothesis by comparing several acoustic properties of /k/ in various vowel contexts with those of /t͡s, t͡ʃ, t͡ɕ/ for three languages differing in stop burst aspiration, i.e., German, Polish and Catalan. Results suggest that the acoustic equivalence hypothesis could account for velar softening in aspirated velar stops but not in unaspirated velar stops. The results also provide an explanation as to why aspirated velar stops are prone to undergo softening more easily when followed by front vocalic segments than in other contexts and positions.

1. Introduction

Velar Softening, i.e., the change of a velar stop into /t͡ʃ/ or /t͡s/ is a widely attested process which takes place in various vowel contexts. For instance, according to the First Slavic Velar Palatalization /k/ changed to /t͡ʃ/ before /j/ and front vowels. The subsequent Second Slavic Velar Palatalization converted /k/ to /t͡s/ before short /e/ and short /i/ ([3]). Velar Softening is also frequently found in Romance, Bantu and other language families.

Two explanatory hypotheses have been proposed in order to account for this sound change. First, velar softening has been explained on an articulatory basis by claiming that the sound change is triggered by particular anterior, (alveolo)palatal realizations of the velar stop in several contextual and positional conditions. In fact, the articulation-based hypothesis has been the only available hypothesis for several decades ([1, 9]). The other, more recent hypothesis ([2, 4, 6]) has its roots in acoustics and claims that velar softening is associated with acoustic equivalence between the spectral cues for a front velar stop, i.e., a velar stop before a front vowel or glide, and the resulting affricate. In these circumstances, the stop burst shows a long /ʃ/-like frication with a 2500-3500 Hz spectral peak, and both front /k/ and /t͡ʃ/ are accompanied by rising F2 vowel transitions.

Production and perception data for Majorcan Catalan ([7, 8]) suggest that velar softening for unaspirated velar stops has an articulatory origin. CV stimuli containing the burst of the unaspirated (alveolo)palatal stop allophone [c] of underlying /k/ and some portion of the following vowel could be identified as /t͡ʃ/ by Catalan informants, more often when the stop preceded a low vowel than when it preceded a front vowel, and fairly often when the stop occurred word-finally. This result may account for the replacement of /k/ by /t͡ʃ/ via [c] in the same conditions in Romance dialects and other language families where [c] is present either as a phoneme or as an allophone of /k/.

Another study ([4]) has shown that analogous stimuli with English aspirated front velar stops may be heard as /t͡ʃ/ when mixed with high intensity white noise. A reason why /k/ before /j, i, e/ was better perceived as an affricate when aspirated than when unaspirated may be because aspirated stops exhibit longer and perhaps more intense bursts than unaspirated stops. A possible problem with Guion's study is that the presence of too low a signal-to-noise ratio in her stimuli could have caused a significant increase in the number of /t͡ʃ/ identification responses, i.e., the acoustic-equivalence hypothesis worked out to some extent only if the white noise was overimposed on the stimuli. The two above studies agree that an increase in the acoustic prominence of the stop burst contributes to higher affricate identification percentages.

Within this framework, this paper seeks to shed light onto the acoustic factors causing front /k/ to be confused with /t͡ʃ/ and other affricates in languages and dialects lacking [c] and exhibiting different degrees of stop burst aspiration:

(i) German, where velar stops are strongly aspirated,
(ii) Polish, where they are slightly aspirated, and
(iii) Eastern Catalan, where they are unaspirated.

The main hypothesis is that strongly aspirated front /k/ should be more prone to be heard as /t͡ʃ/ than moderately unaspirated front /k/ while unaspirated front /k/ should yield essentially no /t͡ʃ/ responses.

2. Methodology

Data include the sequences /kj, ki, kɛ, ka, kɔ, ku, t͡sa, t͡ʃa/ and, in addition, /t͡ɕa/ for Polish as a possible palatalization output. Affricates may exhibit language-dependent differences in place of articulation: /t͡s/ is dental in Polish, dentoalveolar in German and centroalveolar or even postalveolar in Catalan; /t͡ʃ/ is alveolopalatal or palatoalveolar in Catalan and German, and apico-postalveolar in Polish; Polish /t͡ɕ/ is alveolopalatal.

It is predicted that the burst acoustic prominence will be higher for /kj, ki, kɛ/ than for the other sequences and that, judging from the frequency of application of velar softening in synchronic and diachronic processes, it will be higher for /kj/ than for /ki/ and for /ki/ than for /kɛ/ ([5]).

2 German, 2 Polish and 2 Catalan speakers read a set of controlled, short sentences with bisyllabic words including the sequences of interest in stressed word-initial position. The words were inserted in the following sentences: (i) German: *Ich habe __ gesagt* (7 syllables) 'I have __ said'; (ii) Polish: *Powiedziala __ do niego* (9 syllables) 'She said __ to him'; (iii) Catalan: *Ell deia __ fort* (6 syllables) 'He was saying __ loud'. Recordings were performed at 44 kHz and downsampled to 20 kHz for analysis.

Several acoustic cues for stops were submitted to analysis:

(i) Frequency of the front-cavity dependent burst spectral peak on autocorrelation LPC spectra using a 25 ms window and 14 coefficients. The measurement point was set at the onset of the frication noise immediately after the burst spike.

(ii) VOT duration, which was obtained by adding up the durations of the burst spike, the following burst frication period, and the burst aspiration period wherever available.

(iii) Burst absolute intensity, as well as burst relative intensity with respect to the intensity level of the adjacent vowel. Intensity values were taken at the same frame selected for measuring the spectral peak frequency.

(iv) Endpoints and ranges of the F2 and F3 vowel transitions. Formant frequency measurements were taken in 10 ms steps placing a cursor in the middle of the formants, and on LPC spectra if spectrographic readings were judged not to be reliable enough. The frequency ranges of the vowel formant transitions were computed by subtracting the frequency values at the endpoint of the formant transitions from those obtained at the midpoint of the vowel.

As for affricates, measures included the spectral peak frequency and intensity of the frication noise, duration of the burst spike + frication noise, as well as the endpoints and ranges of the vowel transitions.

ANOVAs with repeated measures were performed on the measures (i) through (iv) for the /kV/ sequences across languages with 'vocalic segment' as the only factor. The factor levels were 'j', 'i', 'ɛ', 'a', 'ɔ', 'u'. Each of the six speakers under analysis contributed one averaged score per condition. Huynh-Feldt corrected degrees of freedom were applied and Bonferroni post-hoc tests were run on the significant effects. The degree of significance was set at $p < 0.05$.

3. Results

3.1 VOT

As shown in Figure 1, VOT durations for /k/ vary in the progression German (80-115 ms) > Polish (35-65 ms) > Catalan (25-65 ms), and are significantly higher for /j, i/ than for non-high vowels ($F(3.7, 18.6)=15.76$, $p =0.000$).

The duration of the frication period for all affricates is about 85-125 ms, except for Polish /t͡ʃ/ which shows a short 50 ms frication period.

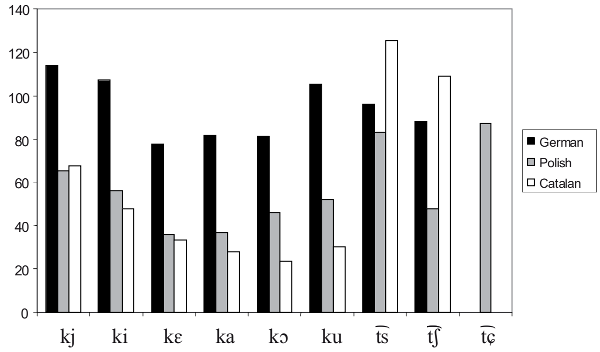

Figure 1: *VOT (in ms)*

3.2 Burst spectral peak

The burst spectral peak frequency is comparable across languages and yielded highly significant contextual effects (F(5, 25)= 188.00, p= 0.000). According to results from post-hoc tests, contextual differences decrease significantly in the following progression (see Figure 2): /j, i/ (2850-3250 Hz) > /ɛ/ (1900-2350 Hz) > /a/ (1600-1800 Hz) > /ɔ/ (1050-1150 Hz) > /u/ (750-900 Hz).

The burst peak frequency for /t͡ʃ/ (2600-3500 Hz) resembles that of /kj, ki/, while that for /t͡s/ is located at a much higher frequency in German and Polish (6500-7500 Hz) than in Catalan (3750 Hz). The Polish affricate /t͡ɕ/ has an intermediate burst spectral frequency between that for the two other affricates (4000 Hz).

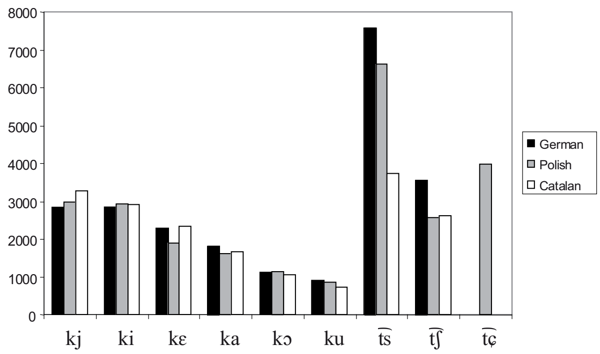

Figure 2: *Burst spectral peak (in Hz)*

3.3 Transition endpoint

Analogously to the burst spectral peak, the F2 transition endpoint frequencies are comparable across languages and yielded significant contextual effects (F(2, 10.2)=86.56, p= 0.000). According to results from post-hoc tests, contextual differences vary for /j, i/ (1950-2250 Hz) > /ɛ/ (1800-1950 Hz) > /a/ (1300-1600 Hz) > /ɔ, u/ (750-1150). Regarding the F3 vowel transition endpoints, significant results were obtained for /j/ and, less so, /i/ (2700-3200 Hz) vs the other vowels (2100-2500 Hz) (F(3.7, 18.6)=18.79, p= 0.000).

The F2 transition endpoints for the affricates resemble those of /k/ in the context of /a/ (/t͡ʃ/ = 1350-1600 Hz; /t͡s/ = 1350-1500 Hz) while those for /t͡ɕ/ are higher (1850 Hz). As for the F3 transition endpoints, the affricate shows intermediate values between /kj, ki/ and /ka/ (/t͡ʃ, t͡s/ = 2250-2600 Hz, /t͡ɕ/ = 2550 Hz).

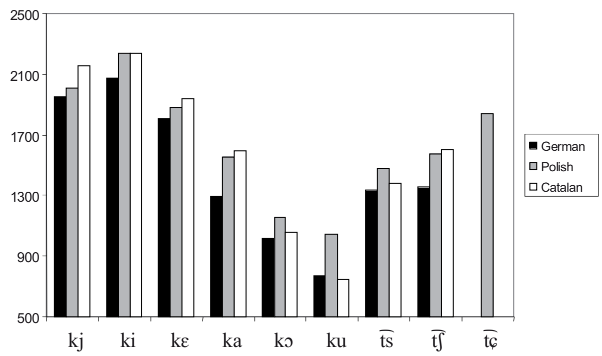

Figure 3: *F2 transition endpoint (in Hz)*

3.4 Transition range

F2 transition ranges turned out to be significantly larger in the context of /j/ (450-675 Hz) than in those of all other vocalic segments (-100/+300 Hz) (F(2.9, 14.6)= 23.69, p= 0.000). They were also large and positive for /ɛ, a/ at least in Catalan. F3 transition ranges were also significantly larger in the context of /j/ (3.2, 16)= 9.49, p= 0.001).

Positive F2 transition ranges for affricates next to /a/ were less prominent than those for /kj/ and more similar to those for /ki, kɛ/, i.e., 0-225 Hz.

3.5 Intensity

Context-dependent effects on absolute and relative stop burst intensity turned out to be non-significant, though mean values happen to be somewhat higher for non-high (55-65 dB) compared to high (52-60 dB) vocalic segments. These contextual differences are in agreement with differences in front cavity size.

Values for /t͡ʃ/ and, less so, for /t͡s/ are similar to the highest intensity values for /k/ in the context of /a/ in German and Polish, i.e., about 55-65 dB (absolute) and 0.70-0.80 (relative). In Catalan the affricate is more intense than the stop, i.e., 65 dB (absolute) and 0.95 (relative).

4. Discussion and Conclusions

Based on VOT values for /k/ and affricates reported in the Results section, it appears that, at least at the production level, German aspirated front velar stops are more prone to be integrated as affricates than their slightly aspirated or unaspirated cognates in Polish and Catalan. Burst intensity level is also similar for both stops and affricates in German. Moreover, frequency data for the burst spectral peak and the transition endpoints and ranges show contextual differences in the progression front > low > back rounded, thus indicating that the replacement of /k/ by /t͡ʃ/ ought to take place mostly before /j/ and more so before /i/ than before /ɛ/. Long-range vowel transitions for the sequence /kj/ are associated with the gliding nature of /j/.

In principle, these findings would be in support of the acoustic equivalence hypothesis stating that aspirated velar stops ought to change to affricates due to acoustic similarity between the input and output segments. The acoustic equivalence hypothesis could also provide an explanation as to why aspirated velar stops are prone to undergo softening more easily when followed by /j, i, ɛ/ than by /a/, schwa or word finally. This hypothesis needs to be tested, however. The perceptual effectiveness of these velar stop spectral, duration and intensity properties in affricate identification will be ascertained in the future by means of perceptual tests.

The articulatorily based hypothesis, on the other hand, would apply to unaspirated velar stops which are expected to yield /t͡ʃ/ via the (alveolo)palatal stop [c]. Accordingly, a preliminary perceptual evaluation of the sequences submitted to analysis in the present paper reveals that Eastern Catalan unaspirated front velar stops cannot be identified with affricates, which leads us to hypothesize that velar softening can only be triggered by (alveolo)palatal stop realizations in this case. This would also answer the question why softening of unaspirated velar stops may take place not only before front vowels and glides but also before low and central vowels as well as word-finally. Burst intensity and formant frequency ranges appear to favour the process implementation in the latter contextual and positional conditions.

5. Acknowledgements

This research has been supported by the projects HUM2006-03742 of the Ministry of Education and Science of Spain and FEDER, 2005SGR864 of the Catalan Government, and GWZ-4/11-1-P2 of the German Research Foundation.

6. References

[1] R. Anttila, *An Introduction to Historical and Comparative Linguistics*. MacMillan, New York, 1972.

[2] S. Chang, M.C. Plauché & J.J. Ohala. Markedness and consonant confusion asymmetries. In: E. Hume & K. Johnson (eds). *The Role of Speech Perception in Phonology*. Academic Press, New York, 2001: 79-101.

[3] T.R. Carlton. *Introduction to the Phonological History of the Slavic Languages*. Slavica Publishers, Columbus, Ohio 1991 .

[4] S.G. Guion. The role of perception in the sound change of velar palatalization. *Phonetica*, 55: 18-52, 1998.

[5] T.A. Hall, S. Hamann & M. Zygis. The phonetic motivation for phonological stop assibilation. *Journal of the International Phonetic Association,* 36: 59-81, 2006.

[6] J.J. Ohala. What's cognitive, what's not, in sound change. In: G. Kellermann & D. Morrisey (eds.). *Diachrony within Synchrony: Language History and Cognition*. Peter Lang Verlag, Frankfurt, 1992: 309-35.

[7] D. Recasens & A. Espinosa. Articulatory motivation of velar softening. In: M. Solé, D. Recasens & J. Romero (eds.). *Proceedings of the 15th ICPhS*. Causal Productions, Barcelona, 2003: vol. 2, 1903-1906.

[8] D. Recasens & A. Espinosa (submitted). Acoustics and perception of velar softening for unaspirated stops.

[9] J.-P. Rousselot. *Principes de phonétique expérimentale*. Leipzig: Welter, 1924-25.