PHONETIC VARIATION OF WORD-FINAL AFFRICATES IN ROMANIAN CONNECTED SPEECH

Oana Niculescu[1] & Andrei A. Avram[2]

[1]Romanian Academy Institute of Linguistics "Iorgu Iordan – Al. Rosetti", Bucharest, Romania
[2]University of Bucharest, Department of English, Romania
oeniculescu@yahoo.com; andrei.avram@lls.unibuc.ro

ABSTRACT

Studies dealing with connected speech phenomena on Romanian data are rather scarce, mainly because of the lack of available corpora suited for analyses at the interface between phonetics and phonology. Consequently, we investigate an understudied topic related to phonetic variation of word-final affricates based on a recently developed open-access standard Romanian speech corpus.

A total of 686 tokens were manually extracted from 4hs of speech pertaining to 4 native speakers (2 female, 2 male). In terms of reduction processes, our results show that the affricates undergo coarticulation in 52% of the data. Fricativization is the preferred outcome (78%), followed by total assimilation (13%), deletion (8%), and (de)voicing (0.5% each), with male speakers showing a higher degree of coarticulation than female speakers ($\chi^2(1)$ = 16.1, $p < 0.001$).

Keywords: Romanian affricates, phonetic variation, Romanian spontaneous speech, corpus linguistics.

1. INTRODUCTION

Among European languages, Romanian is classified as an under-resourced language [1]. In terms of the few topics explored having a speech technology focus, we would like to mention that of sound change in synchronic variation [2], lenition and fortition of stop codas [3], as well as the deletion of the definite article *-l* [4] and vowel hiatus reduction processes in controlled and spontaneous speech [5, 6]. Nevertheless, studies dealing with connected speech phenomena on Romanian data are still rare.

Through this research paper we aim to fill in this gap by investigating an understudied topic related to phonetic variation of word-final affricates. The analysis is carried out on a recently developed open-access standard Romanian speech corpus. We will be examining coarticulation processes conditioned by the phonological context, segment duration, voicing degree, as well as the gender of the speakers.

Our analysis centres around the three affricates present in contemporary standard Romanian:

o   voiceless dental /ts/ (orthographic *ţ*)

baţi 'to beat'   /batsj/ [batsʲ]
o   voiceless palatal /tʃ/ (orthographic *c + e/i*)
baci 'shepherd'   /batʃj/ [batʃʲ]
o   voiced palatal /dʒ/ (orthographic *g + e/i*)
bagi 'put in₂SG'   /badʒj/ [badʒʲ]

The voiced dental affricate /dz/ appears only in certain Daco-Romanian subdialects, being spread mainly in the NW (Maramureș) and SW (Banat) areas of the country [see 7, 8].

It is important to point out that, especially in word-final position, Romanian is the only Romance language to display secondary palatalization [5, p.545]. All consonants can be affected by it. In this context, Romanian /i/ does not always surface as a vowel. Under certain morpho-phonological circumstances, such as plural forms of nouns, adjectives, and certain verb forms (*face* [fatʃe] 'do₃SG' – *faci* [fatʃʲ] 'do₂SG'), /i/ is reduced to a palatal gesture [ʲ] [9: 537]. This morpheme, in turn, can trigger a change in PoA, assibilation or coronalization in the preceding alveolar or velar obstruent (*vag* [vag] 'vagueADJ.M.SG' – *vagi* [vadʒʲ] 'vagueADJ.M./F.PL') [10]. Among the three affricates, only the voiceless dental /ts/ can occur in absolute word-final position (*hoţ* [hots] 'thief' – *hoţi* [hotsʲ] 'thiefPL'). Even though we focus on the subset of word-final affricates, these consonants can also appear in initial and medial position in the word (for a phonotactic description, see [11]).

The first studies dealing with phonetic reduction of word-final affricates in Romanian dialectal texts revealed that fricativization is the most frequent phenomena (307 cases), followed by deletion (60 cases) and (de)voicing (16 cases) [12]. The results also showed that the voiced palatal affricate did not undergo devoicing or deletion [12: 102]. The analyses later continued with pioneer acoustic studies carried out on a Kay Sonagraph investigating the overall duration of the palatal [13, 14] and dental affricates [15] in contrast to the duration of the corresponding plosive + fricative cluster duration. The data were derived from monosyllabic words ("logatoms") with the affricates in intervocalic position. The clusters scored an average duration of 270ms, 261ms and 212ms, respectively, while the affricates had a mean duration of 191ms /tʃ/, 185ms /ts/, and 150ms /dʒ/.

This present study complements previous research on the topic by extending the analysis toward

connected speech data. To our knowledge, this would be the first phonetic account of word-final Romanian affricates documented on naturalistic data.

2. DATA AND METHODOLOGY

2.1. Naturalistic data

This analysis is based on 4 hours of addressed monologues pertaining to 4 adult native speakers, 2 male (136 minutes) and 2 female (17 minutes), ages 33 to 37 (mean = 35.3, SD = 1.7). Participants share the same socio-economic background and are representative of the Southern dialect on which the standard language is based. The data are extracted from a larger Romanian speech corpus (12 participants, 6M/6F), *Ro-Phon*, developed and transcribed by the main author as part of her postdoctoral research project (2020 – 2022) [16]. The corpus was designed for eliciting both spontaneous speech (through a monologue task), as well as controlled speech (by means of a reading task, token words placed in carrier sentences), with recordings carried out in a sound-attenuated room, using a stand microphone connected to a laptop via an external audio interface. All recordings on connected speech data share the same three conversational topics related to past activities (memories from childhood, life lessons), present pursuits (activities, likes and dislikes, travelling), and future projects (personal as well as professional), making the data comparable. Both the audio files and the corresponding TextGrids containing the ortho- graphic and phonological transcriptions are available, alongside the necessary metadata for each speaker.

2.2. Variables studied

We focus our analysis on word-final affricate reduction processes in connected speech related to:

o fricativization /fat͡ʃj kum/ [faʃ.kum]
/nit͡ʃj məkar/ [niʒ.mə.kar]
o total assimilation /pot͡sj sə/ [pos.sə]–[po.sə]
o deletion /alt͡sj ʒukətorj/ [al.ʒu.kə.torj]
o voicing /fat͡ʃj via/ [fad͡ʒ.vi.ja]
o devoicing /aʒund͡ʒj#/ [a.ʒunt͡ʃ]

It is important to note that these coarticulatory processes are preceded by the deletion of the word-final glide.

In relation to the phonological context, so as to make the results comparable with other studies on Romance data [17], we sorted the tokens into 5 categories depending on the following phonological context, namely when the next word begins with a: *voiceless obstruent* /invət͡s pe/, *voiced obstruent* /pot͡sj da/, *sonorant* /nit͡ʃj nu/, *vowel* /merd͡ʒj in/, or *pause* (silence or breath) /mind͡ʒj#/.
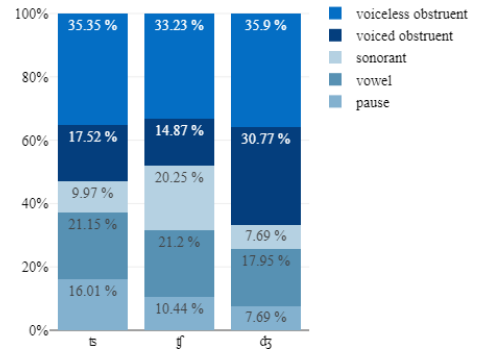
Figure 1: Percentage of word-final affricates (in group) as a function of the following context

We observe that the data are not evenly distributed across the five contexts, with voiceless obstruents having a frequency of 34% (N = 236), vowels 21% (N = 144), voiced obstruents 17% (N = 117), sonorants 15% (N = 100), and pause 13% (N = 89). Based on the phonological context, we expect to find a lower frequency of coarticulatory processes when the affricates are followed by a vowel, as opposed to a higher degree of coarticulation before word-initial obstruents.

For measuring the proportion of voicing in the constriction, we used the Voice Report in Praat [18] showcasing the fraction of locally unvoiced frames. Since the voicing ratios displayed a gradient pattern along the acoustic continuum cuing the voice contrast (Fig. 2), we opted for a ternary classification of voicing degree (similar to [19]): *voiced outputs* (when more than 90% of the interval was identified as voiced by the Voice Report), *partially voiced outputs* (between 10% and 90% of the interval was voiced), and *unvoiced outputs* (less than 10% of the interval was designated as voiced by the Voice Report in Praat).
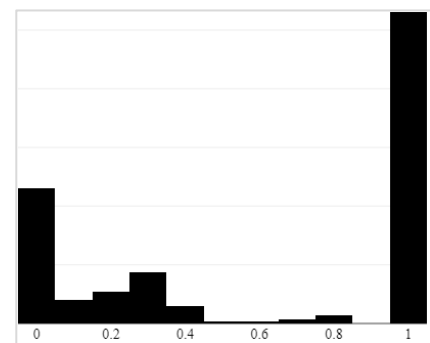
Figure 2: Voicing ratios distribution of word-final affricates undergoing fricativization

The skewness of the data can be, in part, attributed to the higher frequency of voiced segments following the word-final affricates (53%, N = 361), as opposed to 47% (N = 325) corresponding to voiceless

segments. Another explanation could be that all voiced outputs, representing 53% of the data (N = 147), have a voicing degree of 100%, with no tokens present between 90% and 99%. Unvoiced tokens have a frequency of 24% (N = 67), whereas partially voiced outputs are found in 23% of the cases (N = 63).

In relation to duration measurements (i) we predict to find a shorter duration, on average, of the voiced affricate in comparison to the voiceless affricates; (ii) depending on the context, we expect a longer duration of the outputs in pre-pausal context as a result of word-final lengthening, and (iii) in terms of voicing ratios, our predictions are that unvoiced segments will have the longest duration, followed by partially voiced and fully voiced outputs.

In this study we also wanted to test the generalizations related to male and female speech patterns in connection to reduced pronunciation variants characteristic of casual speech. Since all speakers have received higher education and the monologues produced are representative of the standard language, we expect to find an increased percentage of standard pronunciation variants.

3. RESULTS AND DISCUSSION

A total of 686 tokens were manually extracted from 4hs of speech pertaining to 4 Romanian native speakers (2 female, 2 male). The voiceless affricates have the highest frequency (/ts/ 48%, N = 331, with just 25 occurrences in absolute word-final position; /tʃ/ 46%, N = 316), while the voiced affricate /dʒ/ only surfaces in 39 contexts, representing 6% of the data.

In our analysis, affricates undergo coarticulation in 52% of the cases. Fricativization is the preferred outcome (78%), followed by total assimilation (13%), deletion (8%), and (de)voicing (0.5% each).

When maintained, the average duration of an affricate is as follows: /ts/ 166ms (SD = 101), /tʃ/ 152ms (SD = 90), and /dʒ/ 96ms (SD = 50). As expected, the consonants display a shorter average duration in connected speech compared to previous results obtained from read speech data [13, 14, 15].
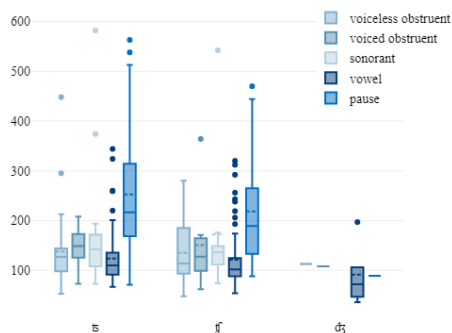
Depending on the phonological context (Fig. 3), we observe a pre-pausal lengthening effect, especially in the case of the voiceless affricates (t-test scores revealed a significant increase in the overall duration of these affricates in the presence of neighbouring pauses, namely for /tʃ/: t = 4.3, df = 36, p < 0.001; for /ts/: t = 2.319, df = 27, p = 0.028). On average, the lowest durations are found when the affricates are followed by a word-initial vowel or voiceless obstruent. The mean duration of each consonant as a function of the following phonological context is presented in Table 1.

	ts	tʃ	dʒ
voiceless obstruent	138 (68)	135 (63)	113
voiced obstruent	149 (40)	151 (93)	108
sonorant	171 (126)	172 (142)	-
vowel	124 (54)	121 (60)	92 (65)
pause	253 (122)	218 (104)	89

Table 1: Mean duration (SD) of affricates as a function of the following context

Word-final affricates undergo fricativization in various degrees. Among the three affricates, /tʃ/ has the highest rate of fricativization (54%, N = 151), followed by /ts/ (38%, N = 105). The voiced palatal affricate undergoes fricativization in only 8% of the cases (N = 21). These observations are in line with previous studied [13, 14, 15].

In terms of phonological context, fricativization occurs especially before an obstruent (39%, N = 107, in the case of voiceless obstruents, and 33%, N = 91, in the case of voiced obstruents). This coarticulation phenomena also occurs, to a lesser degree, before a sonorant (26%, N = 71), a vowel or a pause (1% each).

As predicted, on the basis of duration measurements and degree of vocal fold vibration, the average durations of the resulting fricatives are: unvoiced 71ms (SD = 42), partially voiced 57ms (SD = 23), and voiced 48ms (SD = 13). A t-test indicated

Figure 3: Duration of word-final affricates as a function of the following context
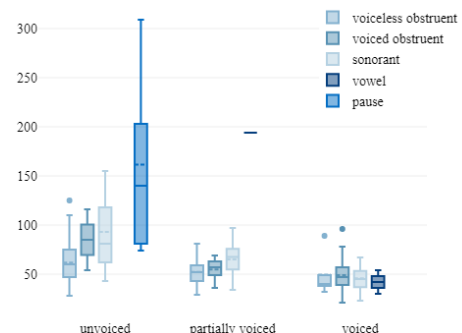
Figure 4: Duration of word-final affricates undergoing fricativization as a function of the following context

a significant decrease in duration from unvoiced outputs to partially voiced (t = 2.392, df = 102, p = 0.018) and, subsequently, voiced fricatives (t = 3.108, df = 81, p = 0.002).

As shown in Fig. 4 (and detailed in Table 2), the highest variation is found in the case of unvoiced fricatives in pre-pausal context (ranging from 74ms to 309ms, with a standard deviation of 98ms). Voiced fricatives in intervocalic position have the shortest duration.

	unvoiced	partially voiced	voiced
voiceless obstruent	62 (21)	52 (12)	50 (23)
voiced obstruent	85 (44)	55 (13)	49 (14)
sonorant	93 (57)	65 (18)	46 (11)
vowel	-	194	42 (17)
pause	161 (98)	-	-

Table 2: Mean duration (SD) of word-final affricates undergoing fricativization as a function of the following context

The second coarticulation process in terms of frequency involving word-final affricates is total assimilation. The process is highly dependent on the following context (see correlation results in Table 3). Out of the three affricates, the dental affricate is frequently targeted (85%, N = 38). The palatal affricates undergo this reduction process to a lesser degree, /tʃ/ 13% and /dʒ/ 2%.

Only 29 tokens exhibit deletion of the word-final affricate, especially before a voiceless obstruent (76%, N = 22). There are only 3 cases in which this reduction process occurs before a sonorant. The voiceless palatal affricate has the highest rate of reduction (66%, N = 19), while /dʒ/ (24%, N = 7) and /ts/ (10%, N = 3) are deleted exclusively when followed by a voiceless obstruent.

Devoicing occurs only in a pre-pausal context, being visible during the frication part of the affricate, while voicing was found exclusively when the voiceless palatal affricate was followed by /v/, resulting in a fully voiced output. We found only two tokens for each of the aforementioned coarticulation phenomena in the corpus. Previous results based on dialectal data [12] indicated a higher rate of deletion, especially targeting /tʃ/, and that both voiceless affricates underwent voicing.

An ANOVA test was used for each coarticulation process in correlation to the phonological context. The results show a strong correlation between the two variables particularly in the case of affricate maintenance (for all of the five contexts) on the one hand, and, on the other hand, in relation to the highest frequent coarticulation processes, namely fricativi-

zation and total assimilation (results summarised in the table below).

	maintenance	fricativization	total assimilation	deletion	voicing	devoicing
voiceless obstruent	77.242 (p < 0.001)	3.685 (p = 0.0553)	105.712 (p < 0.001)	23.797 (p < 0.001)	1.050 (p = 0.3058)	1.050 (p = 0.3058)
voiced obstruent	59.430 (p < 0.001)	92.799 (p < 0.001)	10.018 (p = 0.0016)	0.227 (p = 0.6338)	9.867 (p = 0.0018)	0.411 (p = 0.5214)
sonorant	23.955 (p < 0.001)	48.698 (p < 0.001)	8.294 (p = 0.0041)	0.435 (p = 0.5100)	0.341 (p = 0.5592)	0.341 (p = 0.5592)
vowel	243.414 (p < 0.001)	132.081 (p < 0.001)	13.000 (p = 0.0003)	8.117 (p = 0.0045)	0.532 (p = 0.4661)	0.532 (p = 0.4661)
pause	88.864 (p < 0.001)	55.320 (p < 0.001)	7.234 (p = 0.0073)	4.531 (p = 0.0336)	0.298 (p = 0.5851)	13.684 (p = 0.0002)

Table 3: Correlation (F) values and significance (p) from t-test; the grey cells highlight there is no significant correlation between the variables

When looking at the coarticulated outputs as a function of gender, we observe that male speakers show a higher degree of coarticulation than female speakers ($\chi^2(1) = 16.1$, p = < 0.001). The tendencies for each reduction process (male monologues vs female monologues) can be summarised as follows: 80% (N = 222) – 20% (N = 55) in the case of fricativization, 84% (N = 38) – 16% (N = 7) when the affricates undergo total assimilation, and 76% (N = 22) as opposed to 24% (N = 7) in the case of deletion. However, since the results are dependent on speakers from only one social class and age group, we acknowledge the limitations of our current study. In future socio-phonetic analyses, we consider it relevant to include speakers from various backgrounds.

Summing up, based on our data, we observed that the dental affricate /ts/ undergoes more coarticulation processes than /tʃ/ ($\chi^2(1) = 9.66$, p = 0.002), which, in turn, coarticulates more than /dʒ/ ($\chi^2(1) = 7.69$, p = 0.006).

5. CONCLUSIONS

This study has investigated coarticulatory processes related to word-final affricates in Romanian connected speech based on a recently developed oral corpus.

Analysing the way in which reduction processes occur in connected speech can benefit linguistics and automatic speech recognition models alike. The data acquired through this study can also contribute to the ongoing phonological investigation of fortition and lenition processes occurring in Romance languages.

With respect to future research on the topic, we plan to examine a larger sample of recorded data from the *Ro-Phon* corpus, both in terms of controlled and spontaneous speech, thus enabling an in-depth cross-linguistic analyses of affricates in Romance languages.

6. REFERENCES

[1] Trandabăţ, D., Irimia, E., Mititelu, V. B., Cristea, D., Tufiş, D. 2012. *The Romanian language in the digital age*. In: Rehm, G., Uszkoreit, H. (eds), *META-NET White Paper Studies*. Berlin: Springer, 1–87.

[2] Chitoran, I., Vasilescu, I., Lamel, L., Vieru, B. 2018. Connected speech in Romanian. Exploring sound change through an ASR system. In: Recasens, D., Sánchez Miret, F. (eds), *Production and perception mechanisms of sound change*. Munich: Lincom, 129–143.

[3] Hutin, M., Niculescu, O., Vasilescu, I., Lamel, L., Adda-Decker, M. 2020. Lenition and fortition of stop codas in Romanian. *Proc. SLTU-CCURL* Marseille, 226–234.

[4] Vasilescu, I., Chitoran, I., Vieru, B., Adda-Decker, M., Candea, M., Lamel, L., Niculescu, O. 2019. Studying variation in Romanian: deletion of the definite article -*l* in continuous speech. *Linguistics Vanguard* 5(1), 1–12.

[5] Niculescu, O. 2018. *Hiatul intern şi hiatul extern în limba română contemporană. O analiză acustică* [Internal and external hiatus in contemporary standard Romanian. An acoustic analysis], PhD dissertation, University of Bucharest.

[6] Niculescu, O., Vasilescu, I., Adda-Decker, M. 2019. Duration Patterns of Internal and External Hiatus as a Function of Speaking Style. *Revue roumaine de linguistique* 64 (2), 189–197.

[7] Mărdărescu-Teodorescu, M. 1993. Structura acustică a africatei [d] în graiurile dacoromâne [The acoustic structure of the affricate [d] in Romanian dialects]. *Fonetică şi dialectologie* 12, 113-116.

[8] Rusu, V. 1984. *Tratat de dialectologie română* [Romanian dialectology treatise], Craiova, Scrisul Românesc.

[9] Renwick, M. 2021. Romanian. In: Meisenburg, T., Gabriel, C., Gess, R. (eds) *Manual of Romance Phonetics and Phonology*. Berlin: de Gruyter, 531–558.

[10] Spinu, L. 2010. *Palatalization in Romanian. Experimental and Theoretical Approaches*, PhD dissertation, University of Delaware.

[11] Roceric-Alexandrescu, A. 1968. *Fonostatistica limbii române* [Romanian Phonostatistics], Bucharest, Editura Academiei Republicii Socialiste România.

[12] Mărdărescu, M. 1975. Semioclusivele româneşti în fonetică sintactică (în sandhi) [Romanian affricates in syntactic phonetics (in sandhi)]. *Fonetică şi dialectologie* 9, 99–106.

[13] Mărdărescu, M. 1969. Durata semioclusivelor prepalatale în limba română literară [Duration of prepalatal affricates in standard Romanian]. *Fonetică şi dialectologie* 6, 59–70.

[14] Mărdărescu-Teodorescu, M. 1976. The duration of prepalatal affricates in Romanian. Auditory tests. *Revue roumaine de linguistique* 21 (1), 99–103.

[15] Mărdărescu-Teodorescu, M. 1993. Africata dentală [ţ]. Observaţii acustice [The dental affricate [ţ]. Acoustic observations]. *Fonetică şi dialectologie* 12, 33–43.

[16] Niculescu, O. 2021. Developing linguistic resources for Romanian written and spoken language, In: Rebreja, P., Onofrei, M., Cristea, D., Tufiş, D. (eds) *Proceedings of the 16th International Conference „Linguistic Resources and Tools for Natural Language Processing"*, Editura Universităţii „Alexandru Ioan Cuza", Iași, 21–36.

[17] Hutin, M., Wu, Y., Jatteau, A., Vasilescu, I., Lamel, L., Adda-Decker, M. 2021. Synchronic Fortition in Five Romance Languages? A Large Corpus-Based Study of Word-Initial Devoicing. *Proc. Interspeech 2021*, Brno, 996–1000.

[18] Boersma, P. 2001. Praat, a system for doing phonetics by computer, *Glot International* 5 (9), 341–345.

[19] Davidson, L. 2015. Patterns of voicing in American English voiced obstruents in connected speech. *Proc. 18th ICPhS* Glasgow, paper number 84, https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2015/Papers/ICPHS0084.pdf [accessed 10.11.2022].