Proceedings of the 1st Joint SLTU and CCURL Workshop (SLTU-CCURL 2020), pages 226–234
Language Resources and Evaluation Conference (LREC 2020), Marseille, 11–16 May 2020

Lenition and Fortition of Stop Codas in Romanian

Mathilde Hutin[1], Oana Niculescu[2], Ioana Vasilescu[1], Lori Lamel[1], Martine Adda-Decker[1,3]
[1] Université Paris-Saclay, CNRS, LIMSI, [2] Romanian Academy Institute of Linguistics "Iorgu Iordan - Al. Rosetti",
[3] Université Paris 3 Sorbonne Nouvelle, CNRS, UMR 7018, LPP
[1] Bât. 507, rue du Belvedère, 91405 Orsay, France [2] Calea 13 Septembrie nr. 13, 050711 Bucureşti, Romania
[3] 19 rue des Bernardins, 75005 Paris, France
{hutin, ioana, lori.lamel, martine.adda}@limsi.fr, oeniculescu@yahoo.com

Abstract

The present paper aims at providing a first study of lenition- and fortition-type phenomena in coda position in Romanian, a language that can be considered as less-resourced. Our data show that there are two contexts for devoicing in Romanian: before a voiceless obstruent, which means that there is regressive voicelessness assimilation in the language, and before pause, which means that there is a tendency towards final devoicing proper. The data also show that non-canonical voicing is an instance of voicing assimilation, as it is observed mainly before voiced consonants (voiced obstruents and sonorants alike). Two conclusions can be drawn from our analyses. First, from a phonetic point of view, the two devoicing phenomena exhibit the same behavior regarding place of articulation of the coda, while voicing assimilation displays the reverse tendency. In particular, alveolars, which tend to devoice the most, also voice the least. Second, the two assimilation processes have similarities that could distinguish them from final devoicing as such. Final devoicing seems to be sensitive to speech style and gender of the speaker, while assimilation processes do not. This may indicate that the two kinds of processes are phonologized at two different degrees in the language, final devoicing being more sociolinguistically stigmatized than assimilation.

Keywords: Romanian, lenition, fortition, automatic alignment, pronunciation variant

1. Introduction

A segment is considered as undergoing weakening, also known as lenition, if it is subject to a transformation that ultimately ends with segmental deletion. Conversely, a segment can be said to undergo strengthening, also known as fortition, if it follows the reverse path (even though this definition can be debated; Honeybone, 2008). From the observation of diachronic change (Brandão de Carvalho, 2008), it is known that, when a segment gains a voiced feature, for example, it is lenited, and when it loses it, it is fortified. From a synchronic perspective, consonant weakening and strengthening processes may occur at different degrees in several Western Romance languages (Ryant & Liberman, 2016; Vasilescu & al., 2018 on Spanish; Hualde & Prieto, 2014 on Spanish and Catalan; Hualde & Nadeu, 2011 on Italian; Jatteau et al., 2019a,b on French). Eastern Romance languages, on the other hand, have been less investigated. Chitoran et al. (2015), exploring consonant weakening in Romanian in a corpus of 8 native speakers, reported only rare instances of lenition. Niculescu et al. (submitted) reported very few consonant alternations in Romanian, except for codas, which is not surprising since the coda position is famously prone to neutralization processes cross-linguistically.

This is why the present study will focus on voicing and devoicing, i.e. weakening and strengthening processes, in Romanian stops in coda position. Romanian is indeed a good candidate to investigate variable voicing and devoicing in coda position because it has a regular voicing contrast and robustly allows stops to contrast for voicing in word-final position: e.g. /krap/, "I crack" *vs* /krab/, "crab"; /kot/, "elbow" *vs* /kod/, "code"; /fak/, "I do" *vs* /fag/, "beech".

Since substantial amounts of data are the only way linguists have at their disposal to investigate actual linguistic usage in a statistically significant manner (Coleman et al., 2016), large corpora are a promising means to investigate fine variation phenomena such as final voicing and devoicing. Automatic speech recognition (ASR) systems, traditionally trained on very large amounts of carefully transcribed speech data and written texts, can be used to exploit such corpora for linguistic studies. Yet, such quantities of data and material are available in electronic form mainly in the world's dominant languages (e.g. English, Arabic, Chinese, Spanish, French). Since obtaining large volumes of transcribed audio data remains quite costly and requires a substantial investment in time and money as well as an efficient supervision, developing language technologies for less-resourced languages is less common.

Even though it is spoken by 25 million native speakers around the world, Romanian figures among these less-resourced languages (Trandabăț et al., 2012). As far as we know, corpora of large continuous speech recognition for Romanian are lacking. National oral annotated corpora for Romanian are scarce and rather difficult to access (Mîrzea-Vasile, 2017), needing permission from the coordinator. Either stored on cassette tapes (ROVA, CORV, IVLRA; Dascălu, 2002, 2011; Ionescu-Ruxăndoiu 2002) or magnetic tapes (AFLR; Marin, 1996), there are certain drawbacks which prove harder to overcome such as poor audio quality, vague metadata, unstructured interview making the data challenging to compare, ambiguous policy of data collection and speaker consent. Newer corpora, such as ROMBAC (Ion et al., 2012) or CoRoLa (Barbu Mititelu et al., 2018), are more inclined towards written text data acquisition and processing. During the last decade, however, there have been several attempts to build ASR systems on small corpora (Petrea et al., 2010; Burileanu et al., 2012). As part of the speech technology development in the Quaero program[1], a Romanian ASR system targeting broadcast and web audio was built but the acoustic models were developed in an unsupervised manner similarly to the method employed in Lamel & Vieru (2010), as no detailed annotations were available for the audio training data downloaded from a variety of websites.

A few studies build on these technological advances to investigate linguistic variation and gather information

[1] www.quaero.org

about specific trends of spoken Romanian (Vasilescu et al., 2014, 2019; Renwick et al., 2016; Niculescu et al., submitted). The present study follows this emerging body of literature and the methodology therein. Its aims are threefold: (i) to explore consonant voicing and devoicing patterns in Romanian as instances of lenition and fortition respectively, (ii) to contribute to the overall picture of the advent of lenition and fortition phenomena across languages, and (iii) to gain insight into the Romanian language from large corpora. The main premise is to use automatic alignments of speech data with dictionaries containing specific pronunciation variants to investigate non-canonical realizations of stops in coda position in Romanian. After a brief overview of Romanian phonology in Section 2, Section 3 is devoted to the description of our data and methodology. In Section 4, we tackle the question of non-canonical devoicing, especially investigating the difference between final devoicing and voicelessness assimilation. Section 5 is devoted to the issue of non-canonical voicing. In Section 6, we investigate the sociolinguistic factors behind both lenition and fortition-type phenomena. Finally, in Section 7, we conclude and discuss the results.

2. Romanian Phonology

Romanian is the only surviving Eastern-European Romance language (Rosetti, 1986), descending from the vernacular variant of Latin which branched into Daco-Romanian, spoken north of the Danube, and three south Danubian dialects, Aromanian, Megleno-Romanian, and Istro-Romanian. The northern dialect is what is usually referred to as Romanian, while the southern tongues have the status of oral dialects (Vulpe, 1978: 293). Described as a spoken vernacular until the appearance of the first written texts (Maiden et al., 2013), namely the Letter of Neacşu, dated 1521, Romanian stands out from other Romance languages due to the remarkable unity shown by the north-Danubian subdialects despite external influence.

In terms of vowel inventory, Romanian is unique among Romance languages due to the central vowels /ɨ, ə/ and the two unary diphthongs, /e̯a, o̯a/ (Chitoran, 2002). The consonant system inherited from Latin, i.e. /p, t, k, f, s, b, d, g, v, z/ was enriched in Romanian by /ʧ, ʤ, dz, ʦ, ʃ, ʒ, c, ɟ, h/, without presenting long consonants with phonological function (SOR; Dindelegan, 2016). There are few studies dealing with Romanian phoneme inventory from a statistical viewpoint (see Roceric Alexandrescu (1968) for an in-depth analysis carried out on written texts, and, more recently, Niculescu (2018) for connected speech data).

Romanian is interesting to observe lenition and fortition-type phenomena because, unlike Western Romance languages, Romanian has not undergone lenition diachronically (Brandão de Carvalho & al., 2008; Alkire & Rosen, 2010). It also still displays a voiced/voiceless opposition for all obstruents[2]. Romanian also stands out among Romance languages since it is one of the few, along with French, to include so many codas. From 400 hours of training data and more than 2.5 million tokens[3] (Adda-Decker, 2019), we know that stops in Romanian have the distribution displayed in Figure 1.

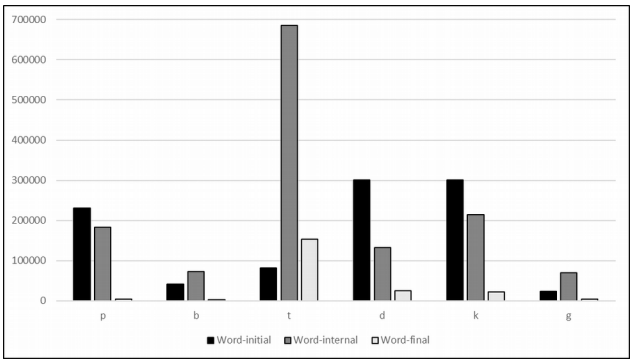

Figure 1. Numbers of stops as a function of position in the word in Romanian.

In this figure, one can see that stops nevertheless tend to appear less in word-final position (8.28%) than in word-initial (38.43%) and word-medial (53.29%) position. Moreover, voiced codas are generally less represented than voiceless ones (26.32% and 73.68% respectively), with a notable dominance, even in coda position, of voiceless alveolars (72.70% of all codas are /t/).

3. Data and Methodology

Consonant alternation in coda position is a very precise issue. Examining this question in large corpora allows the quantification of the variable tendency towards devoicing and voicing under less supervised settings than laboratory recordings, and the larger the corpora, the more precisely the phenomenon can be described (Coleman et al. 2016). Jatteau et al. (2019b) investigated 195,000 items, a substantial amount of data extracted from three large corpora of French. Unfortunately, we do not have access to the same quantity of data for Romanian, which explains why so few large-scale studies on variation have been conducted on this language.

The corpus used for the present study, created by the Quaero program, is representative of Standard Romanian. It is twofold and consists of 3.5 hours of broadcast news, i.e. prepared speech, and 3.5 hours of interviews, i.e. spontaneous speech. More precisely, the first part of the data was gathered from several Romanian radio and television shows (from the RFI Journal and RRA – Radio România Actualități – radio stations and the Euranet news agency) and consists mainly of read and semi-prepared news. Though the number of speakers varies according to the broadcast channel, ranging from 3 to 24, this first part includes a total of 79 different speakers. Broadcasts with significant quantities of overlapping speech and noisy background were excluded. As for the second part, it gathers televised debates recorded from the Romanian national TV channel Antena 3 and includes 50 speakers.

The data have been manually orthographically transcribed and benefited from a speech-to-text alignment with the system described in Vasilescu et al. (2014).

After removing acronymic words followed by the masculine definite marker *-ul* (37 items), the corpus is left

[2] Except for the laryngeal fricative /h/ that can only be voiceless.

[3] Unfortunately, this training data were not manually annotated nor the alignment manually verified, meaning that it is not ideal to observe fine phenomena such as voicing alternation.

with a total of 4529 tokens. Of these, 86% are classified as ending in a canonically voiceless stop, and 14% as ending in a canonically voiced stop. The distribution of these codas is given in Table 1.

Voiceless codas			Voiced codas		
p	t	k	b	d	g
110	3288	486	60	487	98

Table 1. Number of occurrences of each coda.

Furthermore, we build here on the method proposed in Hallé & Adda-Decker (2007) to study voicing alternations through automatic forced alignment introducing specific variants in the pronunciation dictionary. Such pronunciation variants are stored in a lexicon which contains both each word's full (also said canonical) pronunciation and potentially altered (also said non-canonical) variants (Adda-Decker & Lamel, 2017), as shown in Figure 2.

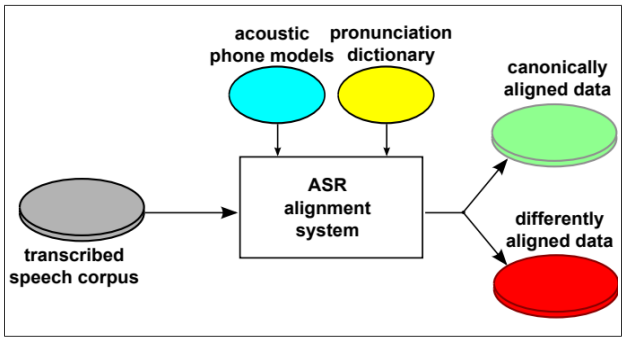

Figure 2. Schema for automatic alignments with pronunciation variants.

A probability is also associated with each variant (Lamel & Gauvain, 2009) and the system will select the most probable variant given the actual acoustic realization. Although they operate categorically and propose only predefined variants, ASR systems offer an alternative method to human perception, which is known to compensate for the missing acoustic information with other available cues (i.e. speech rate, context and word length; Mitterer, 2011). This variant-based approach has given reliable accounts of lenition and fortition-type consonant variation for Romanian before (Vasilescu & al., 2019), as well as for French (Jatteau & al., 2019a,b) and Spanish (Ryant & Liberman, 2016; Vasilescu & al., 2018). During alignment, voicing and devoicing are decided if the best matching phone model corresponds to the voiced or voiceless variant respectively and not to the original canonical voiceless or voiced phone. Hence, for any occurrence of /b, d, g/, its voiceless counterpart may be selected by the system if the acoustic realization of the consonant best matches the corresponding model. For instance, the Romanian word *dialog*, /dialog/ could be transcribed either as [dialog] or as [dialok] depending on whether the system considered the last consonant to correspond to the voiced or voiceless consonant. Conversely, for any occurrence of /p, t, k/, its voiced counterpart may be selected by the system if the acoustic realization of the consonant best matches the corresponding model. For instance, the Romanian word *grup*, /grup/ could be transcribed either as [grup] or as [grub] depending on whether the system considered the last consonant to best correspond to the voiceless or voiced consonant.

This data will allow us to observe devoicing and voicing rates according to the stop's right context, point of articulation, speech style and gender of the speaker.

4. Devoicing Phenomena

Devoicing is a process whereby a canonically voiced consonant such as /b, d, g/ is realized as a devoiced [p, t, k]. It is therefore one manifestation of the wider phenomenon known as fortition.

In the data, 51.78% of /b, d, g/ in coda position are realized as devoiced. Such high rates need to be investigated more finely.

4.1 Voicelessness Assimilation *vs* Final Devoicing

Theoretically, devoicing can be the result of at least two different phenomena: voicelessness assimilation and final devoicing.

Voicelessness assimilation is a process whereby canonically voiced consonants are realized as devoiced due to the presence of a voiceless obstruent in the adjacent left or right context. In the case of codas in synchrony, the assimilation comes from an adjacent right obstruent, such as in French (Snoeren et al., 2006). For instance, Fr. *la soude pue*, "the hold smells bad" can be pronounced /lasu**t**py/ instead of /lasu**d**py/ where canonical /d/ is realized as [t] because of the regressive voicelessness assimilation to [p].

Final devoicing is the process whereby canonical contrastive voiced consonants are devoiced in domain-final position, as in Russian *Youtu[p]*. As pointed out by Jatteau et al. (2019a,b), many factors converge to suggest that final devoicing is a "natural" process. It is widely attested cross-linguistically (Blevins, 2006), it appears regularly in L1 and L2 acquisition (Broselow, 2018), and it constitutes a frequent sound change (Kümmel, 2007; pp. 184-186). Several sources for final devoicing have been proposed in the phonetic literature: lack of the consonant-vowel transition and its cues to the voicing contrast (Steriade, 1999), anticipatory glottal opening for breathing (Myers, 2012), utterance-final decrease of subglottal pressure yielding voicing offset prior to the obstruent release (Westbury & Keating, 1986) and failure of production and perception of voicing in utterance-final lengthening (Blevins, 2006; Ohala, 1997). These phonetic sources predict that variable final devoicing could be found in languages which do not present a phonologized process of final neutralization, such as the one observed in contemporary metropolitan French (Jatteau et al., 2019a,b).

To investigate these two devoicing phenomena, this paper studies the voicing alternations of canonically voiced stops in word-final position in a large corpus of Romanian.

4.2 What Kind of Devoicing in Romanian?

As mentioned above, voicelessness assimilation can be shown to happen if a substantial amount of devoicing appears more before a voiceless obstruent than anywhere else. Conversely, final devoicing proper can be shown to happen if a substantial amount of devoicing appears more before pause than in any other environment.

To investigate this question, the data was broken into five categories: whether the consonant appeared before pause (hesitation, breath or silence) or before a word beginning with a vowel, a sonorant, a voiced obstruent or a voiceless obstruent.

Figure 3 shows the rates of devoiced /b, d, g/ in all five possible contexts. The bars indicate the actual numbers of tokens in the data and the rates proportioned on 100%.

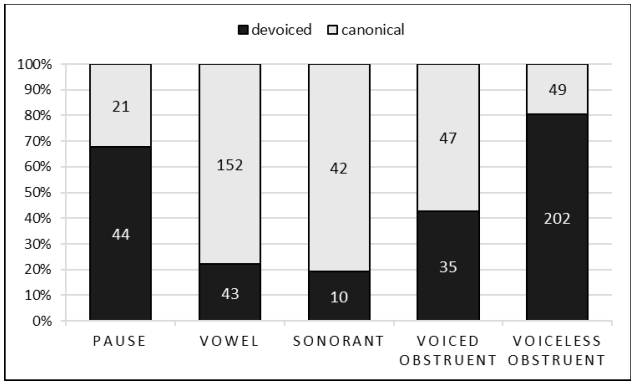

Figure 3: Percentage (and counts) of canonical and devoiced realizations of word-final /b, d, g/ in Romanian as a function of right context.

The data show that there is indeed a high rate of coda devoicing before pause (67.69%), and an even higher rate before voiceless obstruent (80.48%). However, the data also show coda devoicing in contexts where it is not expected, i.e. where devoicing cannot be due to natural devoicing in utterance final position (before pause) nor to assimilation of voiceless segments (before voiceless obstruents). This devoicing before vowel, sonorant or voiced obstruent might be due to the lexical accent, which has not been considered for the present study.

Still, the devoicing rates according to the following context are statistically significant ($\chi^2 = 183.19$, df = 4, p < 2.2e-16). These results indicate that Romanian codas are subject to final devoicing, but most of all are extremely prone to assimilation of the laryngeal feature, in this case of voicelessness.

4.3 Devoicing as a Function of the Coda's Place of Articulation

Since it is more difficult to maintain the pressure differential across the glottis with a smaller vocal tract (Ohala, 1997), we hypothesize that posterior stops will present more devoicing than the anterior ones.

To avoid the bias of the right context and establish the actual part of place of articulation in each type of devoiced codas (the ones undergoing final devoicing *vs* the ones undergoing assimilation), each of them will be investigated separately.

To investigate the role of place of articulation of the coda in final devoicing proper, here we focus on only the canonically voiced codas /b, d, g/ followed by pause (n = 65).

As shown in Figure 4, before pause, the most posterior place of articulation is the one displaying the least final devoicing (63.64%). However, due to the very low number of labial codas in the data, this result is not statistically significant ($\chi^2 = 0.26433$, df = 2, p = 0.8762).

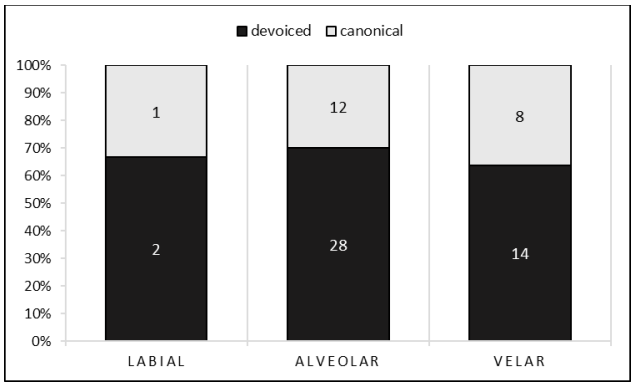

Figure 4: Percentage (and counts) of devoiced codas preceding pause as a function of place of articulation.

To compare those results with the other context favoring devoicing, i.e. the regressive assimilation context, we now focus only on the codas followed by a voiceless obstruent (n = 251), which might provide a better insight on the issue since there are more tokens.

Figure 5 shows the numbers of devoiced codas before voiceless onsets.

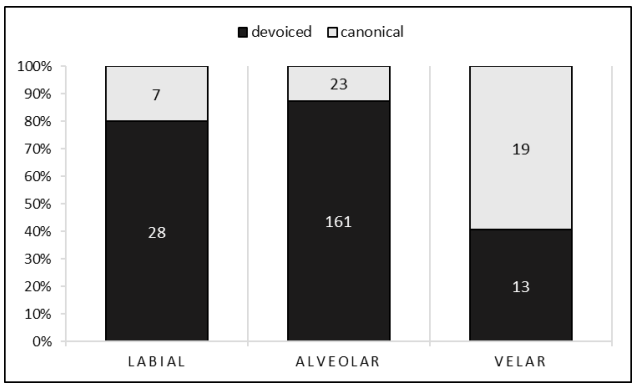

Figure 5: Percentage (and counts) of devoiced codas preceding voiceless obstruent as a function of their place of articulation.

The results here display the same tendencies as for codas before pause: alveolars are the most devoiced (87.50%), followed closely by labials (80.00%), and velars are by far the least devoiced (40.63%). Only in this case, the differences are statistically significant ($\chi^2 = 38.13$, df = 2, p = 5.251e-09).

These results are in line with Jatteau et al. (2019c) who expected more devoicing of velars than of alveolars and labials in French, but found the same results as we do, i.e. more devoicing of labials and alveolars than of velars.

5. Voicing Phenomena

Voicing is the process whereby a canonically voiceless consonant such as /p, t, k/ is realized as a voiced [b, d, g]. It is therefore a manifestation of the wider phenomenon known as lenition.

In the data, 16.47% of /p, t, k/ in coda position are realized as voiced. Although this ratio does not seem as impressive as the one for devoicing, it still represents 639 tokens, which should be enough to gain a valuable insight into voicing patterns in Romanian.

5.1 Regressive Voicing Assimilation

Natural, spontaneous voicing, i.e. in coda position before pause, is controversial, but regressive voicing assimilation, i.e. voicing of codas before voiced obstruent, is well-documented (Snoeren et al. 2006, Hallé & Adda-Decker 2007). This brings us to hypothesize that regressive voicing assimilation can be shown to happen if a substantial amount of voiceless codas are realized as voiced when followed by voiced consonants and not in the other contexts.

To investigate this question, the data was broken into the same five categories: whether the consonant appeared before pause (hesitation, breath or silence) or before a word beginning with a vowel, a sonorant, a voiced obstruent or a voiceless obstruent.

Figure 6 shows the rates of voiced /p, t, k/ in the five possible contexts.

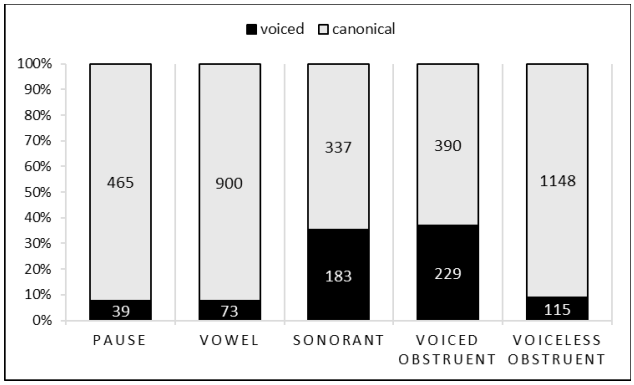

Figure 6: Percentage (and counts) of canonical and voiced realizations of word-final /p, t, k/ in Romanian as a function of right context.

The data show that there are, as expected, rather low rates of voicing before pause (7.74%), vowel (7.50%) or voiceless obstruent (9.11%) and that voicing happens mainly before voiced obstruent (37.00%) or sonorant (35.19%). Since sonorants are voiced by default, they are a possible context for voicing assimilation, and since the differences between our rates are statistically significant (χ^2 = 456.57, df = 4, p < 2.2e-16), we can conclude that non-canonical voicing of codas in Romanian is indeed an instance of regressive voicing assimilation.

5.2 Voicing as a Function of the Coda's Place of Articulation

Since the investigation of place of articulation in devoicing shows unexpected results, it might be interesting to look at the same issue for voicing – especially given the very high number of alveolar voiceless stops in the data.

Figure 7 shows the rates of codas undergoing non-canonical voicing as a function of their place of articulation.

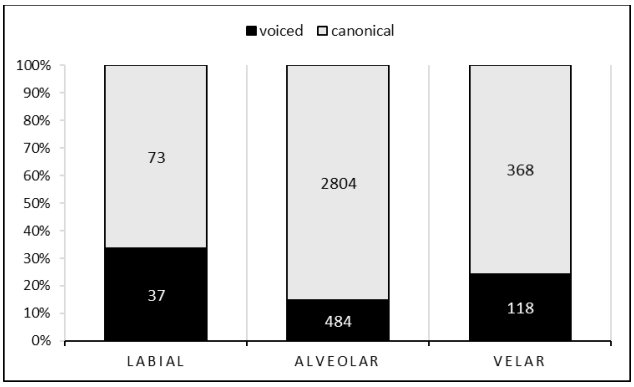

Figure 7: Percentage (and counts) of voiced codas preceding voiced consonants as a function of place of articulation

The results show that labials are more voiced (33.64%) than velars (24.28%) and alveolars (14.72%), the difference between the three places of articulation being statistically significant (χ^2 = 52.471, df = 2, p = 4.036e-12). It is interesting to note that alveolar voiceless stops /t/ are the least voiced when their voiced counterparts /d/ were the most devoiced, and that it could be due to the frequency of /t/ in coda position in Romanian.

6. Sociolinguistic Factors

The data shows that there is a substantial amount of final devoicing proper and of regressive voicelessness and voicing assimilation in Romanian. The strength of such data also resides in the metadata available, which can provide insights into more sociolinguistic questions.

6.1 Non-canonical Realizations Depending on Speech Style

Variation is expected primarily in less formal, spontaneous speech styles (Jatteau et al. 2019a,b; Vasilescu et al. 2019) and we therefore expect more non-canonical realizations in non-prepared rather than in prepared speech. As described in Section 3 about the data and methodology, our data is comprised of two halves: the first can be considered prepared, formal speech in broadcast news (RFI, RRA and Euranet), and the second can be considered as less formal speech in debates and interviews (Antena 3).

For this analysis, we will first consider the effect of speech style on devoicing, then on voicing. In each case, we will only consider consonants in the relevant environment, as to avoid a bias of the right context. We will focus only on canonical /b, d, g/ in contexts of final devoicing (before pause, n = 65) and voicelessness assimilation (before voiceless obstruent, n = 251), and on

canonical /p, t, k/ in contexts of voicing assimilation (before voiced obstruent and sonorant, n = 808).
The results for final devoicing as a function of speech style are displayed in Figure 8.

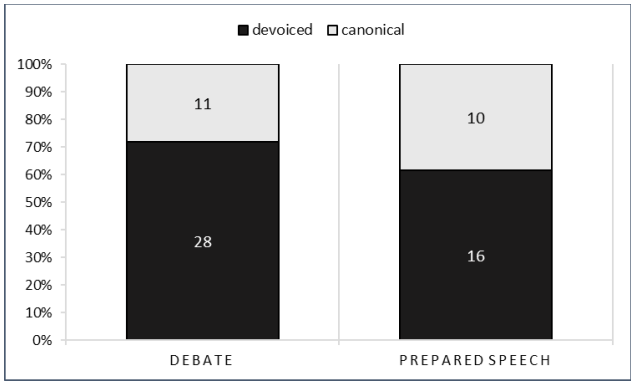

Figure 8: Percentage (and counts) of devoiced codas before pause as a function of speech style.

Figure 8 shows, as expected, more final devoicing in debates (71.79%) than in prepared speech (61.54%). However, the difference between the two speech styles (Δ = 10.26%) is barely significant (χ^2 = 0.35466, df = 1, p = 0.5515).
In the case of voicelessness assimilation however, the results are interesting. There are, overall, more devoiced realizations before voiceless obstruent than before pause, with 82.73% devoicing in debates and 78.72% in prepared speech. However, the difference between speech style is less important (Δ = 4%, χ^2 = 0.40142, df = 1, p = 0.5264). This would indicate that voicelessness assimilation is more generalized across Romanian than final devoicing.
As for voicing, the results of non-canonical realizations of /p, t, k/ before voiced obstruent and sonorant are displayed in Figure 9.

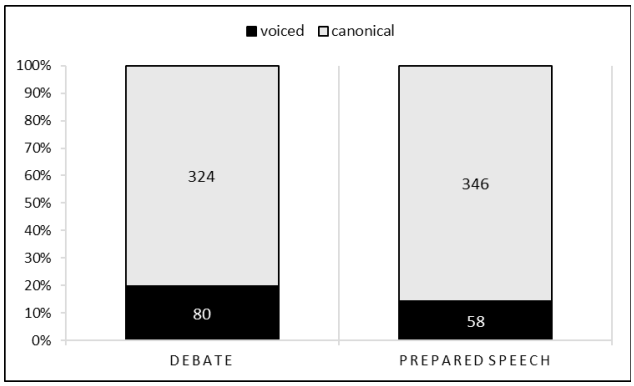

Figure 9: Percentage (and counts) of non-canonically voiced codas preceding voiced consonants as a function of speech style.

Figure 9 shows that there are, as expected, more voicing assimilation in debates (19.80%) than in prepared speech (14.36%), the difference between the two speech styles being statistically significant (χ^2 = 3.8539, df = 1, p = 0.04963).

Overall, then, even though the effect is not significant for final devoicing, non-prepared speech is usually more altered than prepared speech. Interestingly, however, the differences between speech styles are less important for voicelessness assimilation than for voicing assimilation.

6.2 Non-canonical Realizations Depending on Gender of the Speaker

Finally, women tend to avoid speech alterations (Adda-Decker & Lamel 2005). We therefore hypothesize that they would display fewer non-canonical realizations than men.
Again, to avoid the bias of the right context and establish the actual part of gender in final devoicing proper, voicelessness assimilation and voicing assimilation, we focus first only on /b, d, g/ followed by pause (n = 65), then on /b, d, g/ followed by voiceless obstruent (n = 251) and finally on /p, t, k/ followed by voiced obstruents and sonorants (n = 808).
Concerning final devoicing, the numbers displayed in Figure 10 show a higher devoicing rate in male speech (70.00%) than in female speech (60.00%).

Figure 10: Percentage (and counts) of devoiced codas before pause as a function of gender of the speaker.

Once again, the results are not statistically significant (χ^2=0.16942, df=1, p=0.6806). There are too few female tokens in the data to provide a reliable account of the effect of gender on final devoicing in Romanian. However, this result suggests the tendency we might find with more data.
In the case of voicelessness assimilation, there are no effect of gender whatsoever, with 81.08% of devoicing in female speech and 80.23% in male speech. This again shows that regressive assimilation of voicelessness is not comparable to final devoicing in every case: socio-linguistically, it is more widely used, with no difference between speech styles nor genders, which means that it may be less phonologized than final devoicing.
Finally, Figure 11 displays the results for voicing assimilation. Results show that there is less voicing assimilation in female speech (15.33%) than in male speech (17.92%) but the effect is rather slim (χ^2 = 0.66417, df = 1, p = 0.4151).

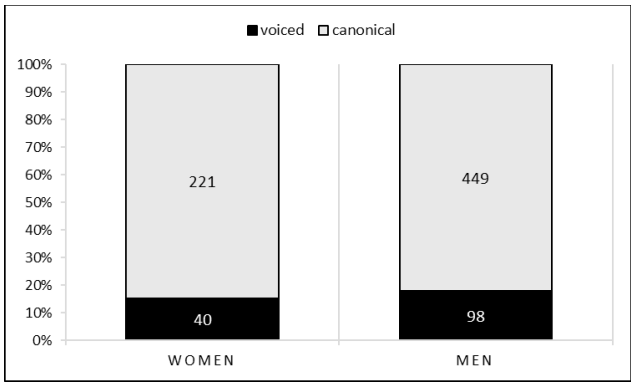

Figure 11: Percentage (and counts) of non-canonically voiced codas preceding voiced consonants as a function of gender of the speaker.

Overall, final devoicing seems to differ more across genders, while assimilation processes seem indifferent to it. This would be an indicator that final devoicing might be more phonologized, to this point, than voicelessness and voicing assimilation in Romanian.

7. Conclusion and Discussion

In conclusion, the present paper aims at providing a first study of lenition- and fortition-type phenomena in coda position in Romanian, a language that can be considered as less-resourced, using several hours of naturalistic speech data.

Our data shows that there are two contexts for devoicing in Romanian: before voiceless obstruent, which means that there is regressive voicelessness assimilation in the language, and before pause, which means that there is a tendency towards final devoicing proper. Data also shows that non-canonical voicing is an instance of voicing assimilation, for it happens mostly before voiced consonants (voiced obstruents and sonorants alike). Although few data were analyzed compared to other Romance languages with stop codas such as French (Jatteau et al. 2019a,b; Hallé & Adda-Decker 2007, Snoeren et al., 2006), trends are consistent with findings in this language, i.e. that there is regressive assimilation of laryngeal feature, that voicelessness assimilates more than voicedness and that final devoicing happens less in velars than hypothesized. However, investigating final devoicing and comparing it to the other two phenomena has proven difficult given the small amount of data available. The limitations of our study show how arduous investigating a less-resourced language can be. To be able to study final devoicing proper, i.e. coda realization before pause, more than 65 tokens in this condition would have been needed, and to study the effect of speech style and gender, more data are needed from more sources and with more varied speakers, especially more women.

Still, two interesting conclusions can be drawn from our analyses. First, from a phonetic point of view, the two devoicing phenomena have the same tendencies regarding place of articulation of the coda, while voicing assimilation displays the reverse tendencies. In particular, alveolars tend to devoice the most, but also to voice the less. Second, an interesting finding is that the two assimilation processes have similarities that could distinguish them from final devoicing as such. Mainly, final devoicing seems sensitive to speech style and gender of the speaker, while assimilation processes do not. This could indicate that the two kinds of processes are phonologized to two different degrees in the language, final devoicing being more sociolinguistically stigmatized than assimilation.

This research is of importance since, although at this stage we don't have enough data to make a decision, previous works point out that including pronunciation variants with model voicing alternation can be helpful (Vasilescu et al., 2018)

8. Acknowledgements

This research was partially supported by the Labex DigiCosme (project ANR-11-LABEX-0045-DIGICOSME) operated by ANR as part of the program "Investissement d'Avenir" Idex Paris-Saclay (ANR-11-IDEX-0003-02).

9. Bibliographical References

Adda-Decker, M. (2006). De la reconnaissance automatique de la parole à l'analyse linguistique des corpus oraux. *Papier présenté aux Journées d'Étude sur la Parole (JEP2006)*, Dinard, France, 12–16 juin

Adda-Decker, M. (2019). Variation in Romance languages: insights from large corpora. *Invited speech at 49th Linguistic Symposium on Romance Languages – LSRL 49*.

Adda-Decker, M. & Lamel, L. (2005) Do speech recognizers prefer female speakers? *Interspeech* 2005: 2205-2208.

Adda-Decker, M. & Lamel, L. (2017). Discovering speech reductions across speaking styles and languages. In Cangemi, F., Clayards M., Niebuhr O., Schuppler B., & Zellers M. *Rethinking reduction: Interdisciplinary perspectives on conditions, mechanisms, and domains for phonetic variation*, De Gruyter Mouton 2017.

Alkire, T. & Rosen, C. (2010). *Romance Languages. A Historical Introduction*, Cambridge University Press, Cambridge.

Barbu Mititelu, V., Tufiş, D. & Irimia, E. (2018). The Reference Corpus of the Contemporary Romanian Language (CoRoLa). *Proceedings of LREC 2018*, Japan, 1178-1185.

Blevins, J. (2006). Theoretical synopsis of Evolutionary Phonology, *Theoretical Linguistics*, vol. 32, 2, 117-166

Brandão de Carvalho, J. (2008). Western Romance. In Brandão de Carvalho, J., Scheer, T. & Ségéral, P. (eds) *Lenition and Fortition*. Berlin: Mouton de Gruyter.

Broselow, E. (2018). Laryngeal contrasts in second language phonology. In L.M. Hymanand & F. Plank, (eds), *Phonological Typology*. Berlin; Boston: De Gruyter Mouton, 312–340.

Burileanu, C., Cucu, H., Besacier, L. & Buzo, A. (2012). ASR domain adaptation methods for low-resourced languages: Application to romanian langage. *European Signal Processing Conference*, 1648– 1652.

Chitoran, I. (2002). *The Phonology of Romanian: A constraint-based approach*. New York, Mouton De Gruyter.

Chitoran, I., Hualde, J. & Niculescu, O. (2015). Gestural undershoot and gestural intrusion – from perceptual errors to historical sound change. *Proceedings of 2nd ERRARE Worskhop* (Sinaia, Romania)

Coleman, J., Renwick, M. E.L. & Temple, R.A.M. (2016). Probabilistic underspecification in nasal place assimilation. *Phonology*, 33(3), 425-458

Dascălu Jinga, L. (2002). *Corpus de română vorbită* (CORV). Eșantioane, București, Oscar Print.

Dascălu Jinga, L. (coord.). (2011). Româna vorbită actuală (ROVA). *Corpus și studii*, Academia Română, Institutul de Lingvistică "Iorgu Iordan – Al. Rosetti".

Dindelegan, G. (ed.). (2016). *The Syntax of Old Romanian*, Oxford, Oxford University Press.

Galliano, S., Geoffrois, E., Mostefa, D., Choukri, K., Bonastre, J.-F. & Gravier, G. (2005). The ESTER phase II evaluation campaign for the rich transcription of French broadcast news. *9th European Conference on Speech Communication and Technology*

Gauvain, J.-L., Lamel, L. & Adda, G. (2002). The LIMSI broadcast news transcription system. *Speech communication* 37 (1-2), 89–108

Gauvain, J.-L., Adda, G., Adda-Decker, M., Allauzen, A., Gendner, V., Lamel, L. & Schwenk, H. (2005). Where Are We in Transcribing French Broadcast News? *Proceedings of ISCA Eurospeech'05*, Lisbon, Sep 2005.

Gravier, G., Adda, G., Paulson, N., Carré, M., Giraudel, A. & Galibert O. (2012). The ETAPE corpus for the evaluation of speech-based TV content processing in the French language. *LREC – 8th International Conference on Language Resources and Evaluation*

Hallé, P. & Adda-Decker, M. (2007).Voicing assimilation in journalistic speech. *16th International Congress of Phonetic Sciences*, 2007, 493–496.

Hallé, P. & Adda-Decker, M. (2011). Voice assimilation in French obstruents: A gradient or a categorical process? *Tones and features: A festschrift for Nick Clements*, De Gruyter, 149–175

Honeybone, P. (2008). Lenition, weakening and consonantal strength: tracing concepts through the history of phonology. In Brandão de Carvalho, J., Scheer, T. & Ségéral, P. (eds) *Lenition and Fortition*. Berlin: Mouton de Gruyter.

Hualde, J. & Nadeu, M. (2011). Lenition and phonemic overlap in Rome Italian. *Phonetica*, 68, 215–242

Hualde, J. & Prieto, P. (2014). Lenition of intervocalic alveolar fricatives in Catalan and Spanish. *Phonetica*, 71, 109–127

Ion, R., Irimia, E., Ştefănescu, D. & Tufiş, D. (2012). ROMBAC: The Romanian Balanced Annotated Corpus. In N. Calzolari, K. Choukri, T. Declerck et al. (ed.), *Proceedings LREC 14,* ELRA, 1235-1239.

Ionescu-Ruxăndoiu, L. (coord.). (2002). Interacţiunea verbală în limba română actuală. Corpus (selectiv). *Schiță de tipologie*, București, Editura Universității din București.

Jatteau, A., Vasilescu, I., Lamel, L., Adda-Decker, M. & Audibert, N. (2019a). "Gra[f]e!" Word-final devoicing of obstruents in Standard French: An acoustic study based on large corpora. In Proceedings of Interspeech, 1726–1730. Graz, Austria.

Jatteau, A., Vasilescu, I., Lamel, L. & Adda-Decker, M. (2019b). Final devoicing of fricatives in French : Studying variation in large corpora with automatic alignment. *International Congress of Phonetic Sciences*, Melbourne, Australie.

Jatteau, A., Vasilescu, I., Lamel, L. & Adda-Decker, M. (2019c). "Gra[f] ! Le dévoisement final dans les grands corpus de français". Paper presented at the SRPP seminar, Université Sorbonne Nouvelle, Feb. 15th 2019

Keating, P., Linker, P. & Huffmann, M.-K. (1983). Patterns in allophone distribution for voiced and voiceless stops. *Journal of Phonetics* 11(3). 277–290.

Maiden, M., Smith, J. C. & Ledgeway, A. (eds.). (2013). *The Cambridge History of the Romance Languages: Structures*, Cambridge, Cambridge University Press.

Marin, M. (1996). Arhiva fonogramică a limbii române (După 40 de ani). *Revista de lingvistică şi ştiinţă literară*, Chişinău, I, 41-46.

Mîrzea-Vasile, C. (2017). Corpusurile de limba română și importanța lor în realizarea de materiale didactice pentru limba română ca limbă străină. *Romanian Studies Today*, I, București, Editura Universității din București, 74-95.

Myers, S. (2012). Final devoicing: Production and perception studies. In T. Borowsky, S. Kawahara, T. Shinya, and M. Sugahara (Eds.), *Prosody matters: Essays in honor of Elisabeth Selkirk*, Equinox, 148-180.

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], Ph.D. dissertation, University of Bucharest.

Niculescu, O., Vasilescu, I., Chitoran, I., Adda-Decker, M. & Lamel, L. (Submitted). Romanian obstruents still strong after all these years: Obstruent voicing and devoicing in a large corpus study of Romanian. *Proceedings of Labphon 17*, 2020.

Ohala, J. J. (1997). Aerodynamics of phonology. *Proceedings of the Seoul International Conference on Linguistics*. Seoul: Linguistic Society of Korea, 92-97

Petrea, C., Ghelmez-Hanes, D., Burileanu, C., Buzo, A. & Cucu, H. (2010). Romanian spoken language resources and annotation for speaker independent spontaneous speech recognition. *Conference on Digital Telecommunications*, 7–10.

Renwick, M.E.L., Vasilescu, I., Dutrey, C., Lamel, L., Vieru, B. (2016). Marginal contrast among Romanian vowels: evidence from ASR and functional load In Proceedings of Interspeech, San Francisco, USA.

Roceric Alexandrescu, A. (1968). *Fonostatistica limbii române*, Bucureşti, Editura Academiei Române.

Rosetti. A. (1986). *Istoria limbii române. I. De la origini până la începutul secolului al XVII-lea*, Bucharest, Editura Ştiinţifică si Enciclopedică.

Ryant, N. & Liberman, M. (2016). Large-scale analysis of Spanish /s/-lenition using audiobooks. *Proceedings of the 22nd International Congress on Acoustics* (Buenos Aires, Argentina)

Snoeren, N., Hallé, P. & Segui, J. (2006). A voice for the voiceless: Production and perception of assimilated stops in French. *Journal of Phonetics*, vol.34, 241–268

Torreira, F., Adda-Decker, M. & Ernestus, M. (2010). The Nijmegen corpus of casual French. *Speech Communication*, vol. 52, no. 3, 201–212.

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

Vasilescu, I., Hernandez, N., Vieru, B. & Lamel, L. (2018). Exploring Temporal Reduction in Dialectal Spanish: A Large-scale Study of Lenition of Voiced Stops and Coda-s. *Proceedings of Interspeech* (Hyderabad, India)

Vasilescu, I., Vieru, B. & Lamel, L. (2014). Exploring Pronunciation Variants for Romanian Speech-to-text Transcription. *Proceedings of SLTU* (St. Petersburg, Russia).

Vasilescu, I., Wu, Y., Jatteau, A., Adda-Decker & Lamel, L. (Submitted). Alternances de voisement et processus de lénition et de fortition : une étude automatisée de grands corpus en cinq langues romanes. *Revue TAL*.

Vulpe, M. (1978). Romanian Dialectology and Sociolinguistics. In A. Rosetti & S. G. Eretescu (eds.), *Revue Roumaine de Linguistique, Current Trends in Romanian Linguistics*, 293-328.

Westbury, J. and Keating, P. (1986). On the naturalness of stop consonant voicing. *Journal of Linguistics*, vol. 22, 145–166.

10. Language Resource References

Quaero Program. (2008-2013). www.quaero.org