



HAL
open science

Ongoing phonologization of word-final voicing alternations in two Romance languages: Romanian and French

Mathilde Hutin, Adèle Jatteau, Ioana Vasilescu, Lori Lamel, Martine Adda-Decker

► To cite this version:

Mathilde Hutin, Adèle Jatteau, Ioana Vasilescu, Lori Lamel, Martine Adda-Decker. Ongoing phonologization of word-final voicing alternations in two Romance languages: Romanian and French. *Interspeech*, Oct 2020, Shanghai (& virtuel), France. 10.21437/Interspeech.2020-1460 . hal-02977812

HAL Id: hal-02977812

<https://hal.science/hal-02977812v1>

Submitted on 26 Oct 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Ongoing phonologization of word-final voicing alternations in two Romance languages: Romanian and French

Mathilde Hutin¹, Adèle Jatteau², Ioana Vasilescu¹, Lori Lamel¹, Martine Adda-Decker^{1,3}

¹ Université Paris-Saclay, CNRS, LIMSI, Bât. 507, rue du Belvédère, 91405 Orsay, France

² Université de Lille, CNRS, UMR 8163, STL, Lille, France

³ Université Paris 3 Sorbonne Nouvelle, CNRS, UMR 7018, LPP, 19 rue des Bernardins, 75005 Paris, France

{mathilde.hutin, ioana.vasilescu, lori.lamel, martine.adda}@limsi.fr,
adele.jatteau@univ-lille.fr

Abstract

Phonologization is a process whereby phonetic substance becomes phonological structure [1]. The process involves at least two steps: (i) a universal phonetic ('automatic') variation becomes a language-specific ('speaker-controlled') pattern, (ii) the language-specific pattern becomes a phonological ('structured') object. This paper will focus on the first step and ask the question of whether three universal phonetic variations of the laryngeal feature of word-final codas (final devoicing, voicelessness assimilation and voicing assimilation) are becoming language-specific patterns in two Romance languages, Romanian and French. Our results suggest that neutralization processes (final devoicing) might be beginning their phonologization process in both French and Romanian whereas assimilation processes (regressive assimilation of voicing and voicelessness) remain universal phonetic tendencies.

Index Terms: Large corpora, forced alignment, phonologization, lenition, fortition, voicing, devoicing, assimilation, French, Romanian

1. Introduction

Relations between patterns of synchronic fine-grained phonetic variation and diachronic change can be investigated today in unprecedented detail thanks to large corpora, automatic language processing technologies and statistical methods. Such data allow large portions of natural speech harvested from real-life conversations to be explored, thus enabling the investigation of very precise phenomena at various stages on their way to (possible) phonologization.

Phonologization is a process whereby phonetic substance becomes phonological structure [1]. The process involves at least two steps: (i) a universal phonetic ('automatic') variation becomes a language-specific ('speaker-controlled') pattern (*phonologization* per se), (ii) the language-specific pattern becomes a phonological ('structured') object (*phonemicization*). During the first stage, a phone has a contextual allophone for anatomic reasons and the allophony becomes (more or less consciously) cognitively controlled by speakers; during the second stage, the allophone progressively gains autonomy until it is used in other contexts than its original one and becomes an independent phoneme.

This paper will focus on the first step and ask the question of whether three universal phonetic variations of the laryngeal feature are becoming language-specific patterns in two

Romance languages, Romanian and French. Metropolitan French word-final codas (henceforth 'codas') have been shown to undergo variable pre-pausal final devoicing [2, 3, 4], voicelessness assimilation and voicing assimilation [5, 6, 7, 8]. So have Romanian codas [8, 9]. The first two phenomena essentially result in (partially) devoiced realizations of canonically voiced stops (i.e. /b, d, g/ pronounced [p, t, k]), while the third one results in the voiced realization of canonically voiceless stops (i.e. /p, t, k/ pronounced [b, d, g]). However, the first one is a case of plain neutralization, i.e. a process whereby a feature (e.g. the laryngeal feature) is lost, while the last two are cases of (regressive) assimilation, i.e. of the laryngeal feature of the following word's onset.

To assess whether these three variational phenomena are becoming phonologized in either or both of these languages, it should be shown that the universal phonetic variation is increasingly more language-specific, i.e. that there are different degrees of variation across linguistic groups. Our hypothesis is that the more a process is phonologized in a language, the more it is cognitively controlled by speakers; and the more it is controlled, the more likely it is to be tagged as a socially marked variable. This means that the more a process is phonologized, the less it is uniformly distributed in the language. To test this hypothesis, we will study final devoicing and regressive assimilation of the laryngeal feature according to two extralinguistic parameters: speech style and speaker gender.

Building on these assumptions, this paper will systematically compare the rates of final devoicing, voicelessness assimilation and voicing assimilation between speech styles and genders so as to uncover in which cases the variation in the realization of the laryngeal feature is uniform and in which ones it is not.

In the remainder of this paper, the Romanian data, the French data and the methodology used for the analysis will be described in Sections 2.1, 2.2 and 2.3 respectively. Section 3.1 is dedicated to the results for final devoicing, 3.2 for voicelessness assimilation and 3.3 for voicing assimilation. Finally, Section 4 concludes and discusses the findings.

2. Data and methodology

Voicing 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 [10].



2.1. Romanian data

The corpus used for the present study, collected under the Quaero program¹, is representative of Standard Romanian [11]. It is twofold and consists of 3.5 hours of broadcast news, i.e. prepared formal speech, and 3.5 hours of interviews, i.e. spontaneous formal 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 comprises 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.

Table 1: Number of occurrences of each coda in Romanian.

	p	t	k	b	d	g	Total
Broadcast	58	1797	241	35	254	40	2425
Interviews	52	1491	245	25	233	58	2104
Total	110	3288	486	60	487	98	4529

2.2. French data

Three corpora are used in the present study. The corpus ESTER [12] originally contains 80 hours of semi-prepared or prepared formal speech (radio broadcast news) but we filtered it to eliminate data from non-metropolitan French (RFI and RTM) and keep only Standard French data. The two-part corpus ETAPE 1 and 2 [13] contains 13.5 hours of radio data and 29 hours of TV data in French, including debates and interviews. Finally, NCCFr [14] is comprised of 31 hours of face-to-face interactions between friends. Each corpus represents a different speech style: prepared formal speech (ESTER), spontaneous formal speech (ETAPE) and casual speech (NCCFr).

When considering only actual phonetic codas, i.e. coda consonants that were not followed by an epenthetic schwa², the corpora are left with 58,893 tokens, as shown in Table 2. Of these, 81% are classified as ending in a canonically voiceless stop, and 19% as ending in a canonically voiced stop.

Table 2: Number of occurrences of each coda in French.

	p	t	k	b	d	g	Total
ESTER	1267	7643	7947	608	2665	325	20455
ETAPE	1488	8620	8112	956	3532	440	23148
NCCFr	423	6850	5624	460	1630	303	15290
Total	3178	23113	21683	2024	7827	1068	58893

2.3. Methodology

Automatic word and phone level alignments of the speech data with their manual orthographic transcriptions were produced using a system derived from that described in [11, 15]. We build here on the method proposed in [6] to study voicing alternations through automatic forced alignment introducing specific variants in the pronunciation dictionary. The pronunciation

¹ www.quaero.org

variants are stored in a lexicon which contains both each word’s full (also called canonical) pronunciation and potentially altered (also called non-canonical) variants [16]. As described in [17], the system selects the most probable variant given the actual acoustic realization.

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 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. The same goes for French /djalɔg/ that could be transcribed either as [djalɔg] or as [djalɔk], and /gʁup/ either as [gʁup] or as [gʁub].

To observe final devoicing and voicelessness assimilation in both languages, we analyzed the realizations of codas before pause (Romanian: n=65; French: n=710) and before voiceless obstruent (Romanian: n=251; French: n=2908) respectively. However, studies have shown that regressive voicing assimilation happens only before voiced obstruents in French [8] but before both voiced obstruents and sonorants in Romanian [8, 9]. We have therefore investigated regressive voicing accordingly (Romanian: n=1139; French: n=8741).

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 [18]). This variant-based approach has given reliable accounts of lenition and fortition-type consonant variation for Romanian before [15], as well as for French [2, 3, 4] and Spanish [19, 20]. This data will allow us to observe neutralization and assimilation rates according to two parameters, speech style and gender of the speaker.

3. Results: degrees of phonologization

Our hypothesis is that universal phonetic tendencies are stable across speaking groups, while phonologizing processes are less so. To test this hypothesis, we analyze final devoicing and regressive assimilation of the laryngeal feature according to two extralinguistic parameters: speech style and speaker gender.

3.1. Final devoicing

Overall, 67.69% of Romanian codas and 48.59% of French codas are devoiced before pause. Do these numbers depend on speech style and / or speaker gender?

3.1.1. Speech style

As mentioned before, our data covers several speech styles: in Romanian, we have data from formal prepared vs formal

² Studies show that word-final schwa tends to disfavor adjacency effects such as final devoicing and regressive assimilation in French [23].



spontaneous speech, while in French we have data from formal prepared (ESTER), formal spontaneous (ETAPE) and informal speech (NCCFr).

As shown in Figure 1, results show that final devoicing is more frequent in debates (71.79%) than in prepared speech (61.54%) in Romanian.

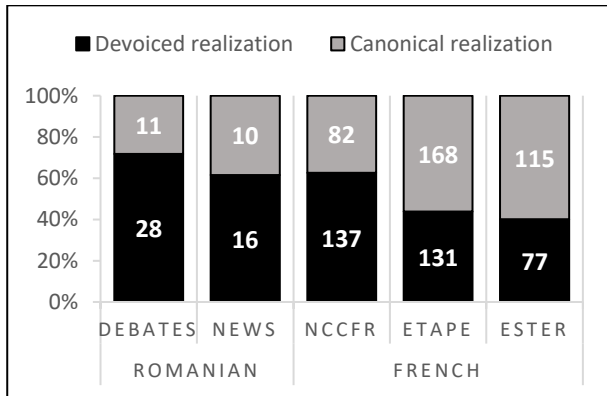


Figure 1: Rates (and counts) of final devoicing as a function of speech style in Romanian and French.

Even though the difference between the two speech styles ($\Delta=10.26\%$) is not significant in Romanian ($\chi^2=0.35466$, $df=1$, $p=0.5$), it shows an interesting tendency similar to that seen for French ($\Delta=22.45\%$ across three speech styles, $\chi^2=25.369$, $df=2$, $p<.0001$). Note however that, in French, the distinction is more between informal and formal speech than between formal spontaneous and formal prepared speech.

3.1.2. Gender of the speaker

Since studies tend to show that there are linguistic differences among genders [21, 22], we hypothesize that both male and female speakers would display various rates of non-canonical realizations, and that the less discrepancy between the two genders, the less phonologized the process.

Final devoicing is subject to a notable difference between genders in both Romanian ($\Delta=10.00\%$, $\chi^2=0.16942$, $df=1$, $p=0.7$) and French ($\Delta=10.52\%$, $\chi^2=6.0135$, $df=1$, $p=0.01$), as shown in Figure 2. However, final devoicing tends to be avoided by women in Romanian but by men in French.

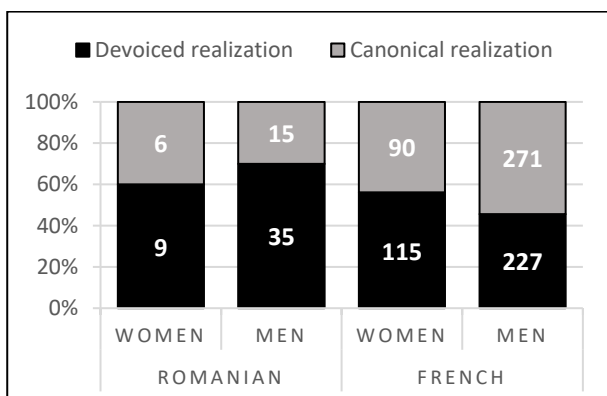


Figure 2: Rates (and counts) of final devoicing as a function of gender of the speaker in Romanian and French.

To sum up, final devoicing is shown to be frequent in both languages, showing variation across speech styles and genders. According to our hypothesis, this would mean that it is

controlled by speakers and therefore potentially on its way to phonologization.

3.2. Voicelessness assimilation

Overall, 80.48% of Romanian codas and 53.51% of French codas are devoiced before a voiceless obstruent.

In the case of Romanian, there are, overall, more devoiced realizations before voiceless obstruent than before pause, with 82.73% devoicing in debates and 78.72% in prepared speech, as shown in Figure 3, which suggests that it is more generalized. Moreover, the difference between speech styles is smaller ($\Delta=4\%$, $\chi^2=0.40142$, $df=1$, $p=0.5$), which indicates that it is more universal. In French, the difference for voicelessness assimilation across speech styles is also less relevant than final devoicing ($\Delta=17.14\%$, $\chi^2=49.946$, $df=2$, $p<.0001$).

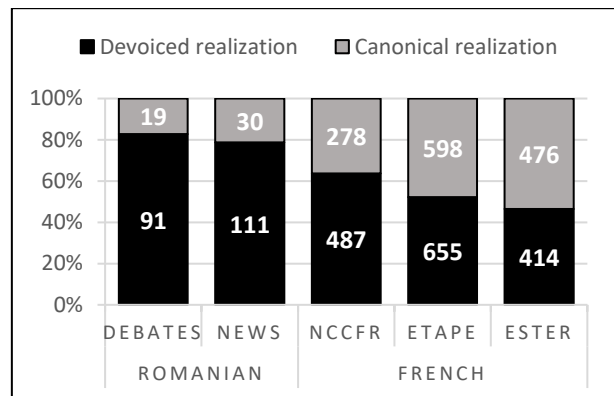


Figure 3: Rates (and counts) of regressive voicelessness assimilation as a function of speech style in Romanian and French.

As for the effect of gender, the difference between genders is very small in both Romanian ($\Delta=0.86\%$, $\chi^2=1.4177e-30$, $df=1$, $p=1$) and French ($\Delta=3.91\%$, $\chi^2=3.4355$, $df=1$, $p=0.06$), as shown in Figure 4.

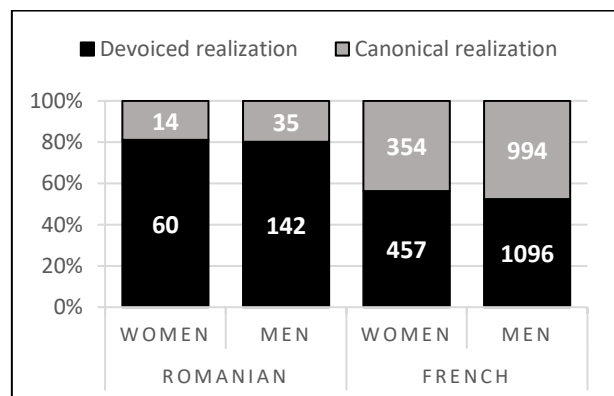


Figure 4: Rates (and counts) of regressive voicelessness assimilation as a function of gender of the speaker in Romanian and French.

To conclude, voicelessness assimilation happens more in French and much more in Romanian than final devoicing, which means it is more generalized. Moreover, there is almost no difference across speech styles or genders, except for the effect of speech style in French (that nevertheless is smaller than that observed for final devoicing), which means voicelessness assimilation not only happens more in both



languages than final devoicing, but is also more uniform across speaking groups.

3.3. Voicing assimilation

Overall, 36.17% of Romanian codas and 76.80% of French codas are voiced as a result of voicing assimilation. It is noticeable that the two devoicing phenomena are more present than the voicing one in Romanian, but the *vice-versa* is observed in French.

As shown in Figure 5, the results of non-canonical realizations of /p, t, k/ before voiced obstruents and sonorants in Romanian show that there is as much voicing assimilation in debates (36.47%) as in prepared speech (35.92%, $\Delta=0.55\%$, $\chi^2=0.016683$, $df=1$, $p=0.9$). As for French, the difference is rather small ($\Delta=5.63\%$) but statistically significant ($\chi^2=93.749$, $df=2$, $p<.0001$). However, the results of non-canonical realizations of /p, t, k/ before voiced obstruents (sonorants not being a context for assimilation in French) are surprising, since the speech style favoring voicing assimilation the most is spontaneous formal ETAPE. As noted in [3, p. 1728], this might be explained “by the more heterogeneous recording conditions of the corpus; portions of dialog overlaps in debates and broadcast interviews may artificially raise the frequency of F0 detection”.

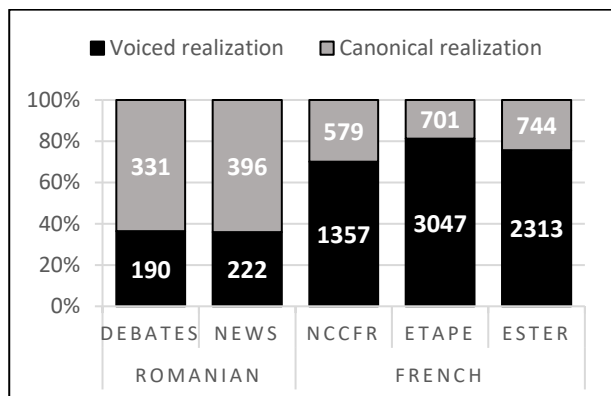


Figure 5: Rates (and counts) of regressive voicing assimilation as a function of speech style in Romanian and French.

The difference for regressive voicing between genders remains rather small but is more noticeable in Romanian ($\Delta=8.87\%$, $\chi^2=8.442$, $df=1$, $p=0.004$) than in French ($\Delta=4.47\%$, $\chi^2=18.64$, $df=1$, $p<.0001$), as shown in Figure 6.

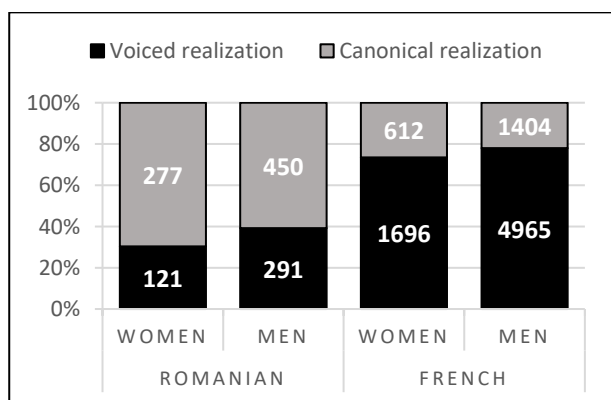


Figure 6: Rates (and counts) of regressive voicing assimilation as a function of speaker gender in Romanian and French.

To summarize, we can say that voicing assimilation is less frequent than both voicelessness assimilation and final devoicing in Romanian but is more frequent in French. However, it is similar to voicelessness assimilation in both languages in that it shows very little variation between speech styles and genders.

4. Discussion and conclusion

This study presented a detailed analysis of one plain neutralization process, final devoicing, and two assimilation processes, voicelessness and voicing assimilation, to investigate their degree of phonologization based on large corpora and ASR systems. The results are summarized in Table 3.

Table 3: Overall rates and differences across groups of speakers (A) in percentage (high rates highlighted in grey)

	Overall rate		Speech Style		Gender	
	RO	FR	RO	FR	RO	FR
Final devoicing	67.69	48.59	10.26	22.45	10.00	10.52
Voicelessness assimilation	80.48	53.51	4.00	17.14	0.86	3.91
Voicing assimilation	36.17	76.80	0.55	5.63	8.87	4.47

As one can see from the overall rates in the table, Romanian favors the two devoicing phenomena over the voicing one, while French favors the voicing phenomenon over the two devoicing ones. However, in both languages, final devoicing is more dependent on speech style and gender than the regressive assimilation processes.

The fact that the tendency is similar in both Romanian and French suggests that both Romance languages are at the same stage of phonologization for non-canonical voicing and devoicing in coda position. It is also possible that voicelessness assimilation is starting the slow process of phonologization in French. Interestingly, however, the two “phonetic” phenomena, which should be “universal”, are present at very different rates in each language. Shedding light on this issue will be a topic for future research.

Further studies should also help establish how advanced non-canonical voicing and devoicing are in both languages in other positions in the word, if Romanian and French are at the same stage of phonologization for other variational patterns, and if other related or non-related languages follow a similar path.

5. 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).

6. References

- [1] L. M. Hyman, “Enlarging the Scope of Phonologization”, *UC Berkeley PhonLab Annual Report*, 4. 2008. Retrieved from <https://escholarship.org/uc/item/3zm91694>
- [2] A. Jatteau, I. Vasilescu, L. Lamel and M. Adda-Decker, “Gra[f] ! Le dévoisement final dans les grands corpus de français”, paper presented at the SRPP seminar, Université Sorbonne Nouvelle, Feb. 15th 2019.
- [3] A. Jatteau, I. Vasilescu, L. Lamel, M. Adda-Decker and N. Audibert, ““Gra[f]e!” Word-final devoicing of obstruents in Standard French: An acoustic study based on large corpora”,



- in *Proceedings of Interspeech*, pp. 1726–1730. Graz, Austria. 2019.
- [4] A. Jatteau, I. Vasilescu, L. Lamel and M. Adda-Decker, “Final devoicing of fricatives in French: Studying variation in large-scale corpora with automatic alignment”, in S. Calhoun, P. Escudero, M. Tabain & P. Warren (eds.), *Proceedings of the 19th International Congress of Phonetic Sciences, Melbourne, Australia*, pp. 295–299, 2019.
- [5] N. Snoeren, P. Hallé and J. Segui, “A voice for the voiceless: Production and perception of assimilated stops in French”, in *Journal of Phonetics*, vol. 34, pp. 241–268, 2006.
- [6] P. Hallé and M. Adda-Decker, “Voicing assimilation in journalistic speech”, in *16th International Congress of Phonetic Sciences*, 2007, pp. 493–496, 2007.
- [7] P. Hallé and M. Adda-Decker, “Voice assimilation in French obstruents: A gradient or a categorical process?”, in *Tones and features: A festschrift for Nick Clements*, De Gruyter, pp. 149–175, 2001.
- [8] M. Hutin, I. Vasilescu, L. Lamel and M. Adda-Decker, “Lénition et fortition des occlusives en codas dans deux langues romanes : le roumain et le français”, in *Actes de JEP 33*. 2020.
- [9] M. Hutin, O. Niculescu, I. Vasilescu, L. Lamel and M. Adda-Decker, “Lenition and fortition of stop codas in Romanian”, in *Proceedings of STLU-CCURL*. 2020.
- [10] J. Coleman, M. E.L. Renwick and R.A.M. Temple, “Probabilistic underspecification in nasal place assimilation”, in *Phonology*, 33(3), pp. 425–458, 2016.
- [11] I. Vasilescu, B. Vieru and L. Lamel. “Exploring Pronunciation Variants for Romanian Speech-to-Text Transcription”, in *Proceedings of SLTU-2014*, pp. 161–168, 2014.
- [12] S. Galliano, E. Geoffrois, D. Mostefa, K. Choukri, J.-F. Bonastre and G. Gravier, “The ESTER phase II evaluation campaign for the rich transcription of French broadcast news”, in *9th European Conference on Speech Communication and Technology*. 2005.
- [13] G. Gravier, G. Adda, N. Paulson, M. Carré, A. Giraudel and O. Galibert, “The ETAPE corpus for the evaluation of speech-based TV content processing in the French language”, in *LREC –8th International Conference on Language Resources and Evaluation*, 2012.
- [14] F. Torreira, M. Adda-Decker and M. Ernestus, “The Nijmegen corpus of casual French”, in *Speech Communication*, vol. 52, no. 3, pp. 201–212, 2010.
- [15] I. Vasilescu, I. Chitoran, B. Vieru, M. Adda-Decker, M. Candea, L. Lamel and O. Niculescu, “Studying variation in Romanian: deletion of the definite article -l in continuous speech”, in *Linguistic Vanguard*, 5/1, 2019.
- [16] M. Adda-Decker and L. Lamel, “Discovering speech reductions across speaking styles and languages”, in F. Cangemi, M. Clayards, O. Niebuhr, B. Schuppler and M. Zellers, *Rethinking reduction: Interdisciplinary perspectives on conditions, mechanisms, and domains for phonetic variation*, De Gruyter Mouton 2017.
- [17] J.-L. Gauvain, G. Adda, M. Adda-Decker, A. Allauzen, V. Gendner, L. Lamel and H. Schwenk, “Where are we in transcribing French broadcast news?”, in *Proceedings of Interspeech 2005 - Eurospeech, 9th European Conference on Speech Communication and Technology*, Lisbon, Portugal, September 4-8, 2005, pp. 1665–1668, 2005.
- [18] H. Mitterer, “Recognizing reduced forms: Different processing mechanisms for similar reductions”, in *Journal of Phonetics* 39(3), pp. 298–303, 2011.
- [19] N. Ryant and M. Liberman, “Large-scale analysis of Spanish /s/-lenition using audiobooks”, in *Proceedings of the 22nd International Congress on Acoustics* (Buenos Aires, Argentina), 2016.
- [20] I. Vasilescu, N. Hernandez, B. Vieru and L. Lamel, “Exploring Temporal Reduction in Dialectal Spanish: A Large-scale Study of Lenition of Voiced Stops and Coda-s”, in *Proceedings of Interspeech* (Hyderabad, India), 2018.
- [21] W. Labov, *Principles of linguistic change, volume 2: social factors*, Oxford: Blackwell. 2001.
- [22] M. Adda-Decker and L. Lamel, “Do speech recognizers prefer female speakers?”, in *Proceedings of Interspeech 2005*: pp. 2205–2208, 2005.
- [23] M. Hutin, A. Jatteau, I. Vasilescu, L. Lamel and M. Adda-Decker, “Le schwa final en français standard est-il un lubrifiant phonétique?”, in *Actes du Congrès Mondial de Linguistique française, CMLF 2020*. To appear.