DURATION PATTERNS OF INTERNAL AND EXTERNAL HIATUS AS A FUNCTION OF SPEAKING STYLE [1]

OANA NICULESCU[2], IOANA VASILESCU[3], MARTINE ADDA-DECKER[4]

Abstract. In this article we investigate duration patterns of internal (adjacent vowels within the word) and external hiatus (adjacent vowels across word boundary) in Romanian as a function of speaking style by comparing data from spontaneous speech and read speech. We look at duration of internal and external hiatus according to primary stress distribution and number of syllables. Results show a systematic reduction of VV-sequences with stress on the first vowels, followed by hiatus with the second vowel stress and unstressed. Heterosyllabic pairs shorten as the number of syllables increases. In total, 53 vowel pairs were analysed (21 word-internal, and 32 word-external). All of the outputs were hand-checked, and the automatic extraction was paired with a manual segmentation.

Keywords: hiatus, automatic speech recognition, Romanian corpus.

1. INTRODUCTION

In this article we aim to provide an up-to-date analysis of hiatus in contemporary standard Romanian, at the interface between phonetics and phonology. In order to get insight into the acoustic patterns that underlie the different realizations of VV-sequences, we compared duration patterns of internal and external hiatus in spontaneous and read speech.

From a phonological standpoint, most of the studies deal with hiatus from an OT (Prince and Smolensky 1993) perspective (Casali 1996 – analysis of VV sequences in 92 languages; Kainada 2007 for hiatus in Greek; Kabak 2007 for hiatus in Turkish; Cabré and Prieto 2006, among others, for hiatus in Spanish; St-Amand 2012 for Québécois French). The only OT account of hiatus resolution in Romanian comes from one of the chapters in Chițoran (2002a). The author describes two ways of avoiding adjacent high vowels (glide epenthesis and glide formation), both related to stress-dependent patterns, and provides a constraint ranking specific to Romanian.

---

[1] This work was partially funded by the Labex EFL mobility grant. The authors are grateful to Ioana Chițoran for insightful comments and discussion.

[2] "Iorgu Iordan–Al. Rosetti" Institute of Linguistics and University of Bucharest, Faculty of Letters, oeniculescu@yahoo.com.

[3] LIMSI, CNRS, Université Paris-Saclay, ioana@limsi.fr.

[4] LIMSI, CNRS, Université Paris-Saclay and Paris 3, madda@limsi.fr.

RRL, **LXIV**, *2*, p. 189–197, Bucureşti, 2019

From an acoustic perspective, analyses focusing on hiatus-diphthong distinction prevail (Aguilar 1999; Hualde and Prieto 2002; Chițoran 2002b; Chioran and Hualde 2007; Alba 2006; Baltazani 2006; Colantoni and Limanni 2008, among others).

For Romanian there are few studies dealing solely with VV-sequences inter- and intra- word (see, more recently, Niculescu 2018a for a monographic account of internal and external hiatus in Romanian; Niculescu 2015 for classification and terminological proposal).

The paper has the following structure: section 2 describes the experimental design; section 3 is devoted to results and discussions; in section 4 we present the conclusions and further studies. We used the following abbreviations: *ExtH* – external hiatus; *Falign* – forced alignment; *IntH* – internal hiatus; *Malign* – manual alignment; *ms* – milliseconds; *n* – number of occurrences; *PrT* – pretonic; *PsT* – post-tonic; *RS* – read speech; *SS* – spontaneous speech; *st.dev* – standard deviation (measured in milliseconds); *V1* – first vowel in the pair; *V2* – second vowel of the pair.

2. EXPERIMENTAL DESIGN

The experiments took place in a quiet room, directly onto a laptop. A Behringer B1 microphone and an external audio interface M-Audio Fast Track were used. In this presentation we focus on data of one Romanian speaker extracted from a larger database of nine subjects representative of the Southern dialect of Muntenia. Each candidate had to perform two tasks. For the first experiment, participants were required to talk about their previous summer activities. For the reading task, all vocalic sequences of IntH and ExtH were extracted from the previous experiment. The tokens were placed in carrier sentences, three repetitions were elicited, and final intonation continuation rise was controlled for. Our selected male speaker produced 40 min of SS and 100 min of RS (35 min for IntH and 65 min for ExtH).

The corpus was recorded, orthographically transcribed and phonetically aligned (manually and automatically) by the first author as part of her PhD research (Niculescu 2018a). For a detailed presentation of the corpus, see Niculescu (2018b).

Based on the waveforms and spectrograms, the manual segmentations carried out in Praat (6.0.37) took place at the following levels:

- *ExtH with pause* (/apə#ɨn/ 'water in' – the duration of the pause varied between 90ms and 4s, with a mean duration of 784ms (st.dev. = 663));
- *ExtH without pause* (/de akum/ 'from now');
- *IntH* (/lua/ 'to take');
- *plurivocalic sequences containing at least one VV-sequence* (/roʃia asta/ 'this tomato').

Our analysis focused solely on word-internal and word-external hiatus (i.e. levels two and three). The duration of the vowel pair was measured starting from the onset of F1 in the first vowel until the offset of F2 in the second vowel.

As a result and a methodological novelty, part of the corpus was forced aligned (one monologue and the entire controlled experiment) with an automatic speech transcription system described in Vasilescu et al. (2014). Annotated corpora have proven to be of high relevance in acoustic analysis, testing various linguistic hypothesis and exploring sound change and variation (Ohala 1996; Adda-Decker 2006). The automatic alignment yielded approximately 45.7k observations which were processed in R. The ASR system has a window length of 30ms, meaning that, when both vowels are aligned, the minimum duration of hiatus is 60ms.

The manual segmentation was paired with the forced extraction, resulting in eight contexts: IntH in RS at Malign and Falign, IntH in SS at Malign and Falign, respectively ExtH in RS at Malign and Falign, and ExtH in SS at Malign and Falign. The total number of vocalic pairs analysed in the experiment was 53 (1244 tokens), out of which 21 IntH (420 tokens), and 32 ExtH (824 tokens).

3. RESULTS

This study investigates duration patterns of hiatus with the aim of highlighting the differences between canonical (phonological) representations of VV-sequences vs. phonetic actualization in continuous speech as portrayed by two speaking styles in Romanian. It is commonly known that connected speech entails systematic variability with respect to various reduction processes (Ernestus and Warner 2011), meaning that hiatus is shorter in spontaneous speech than in careful speech.

Figure 1 presents the overall duration of IntH vs. ExtH in read and spontaneous speech with results from forced and manual alignment. The mean value of each alignment is marked with a white asterisk positioned inside the boxplot, close to the median line. The grey box at the bottom represents the mean duration between Falign and Malign depending on the style of speech.

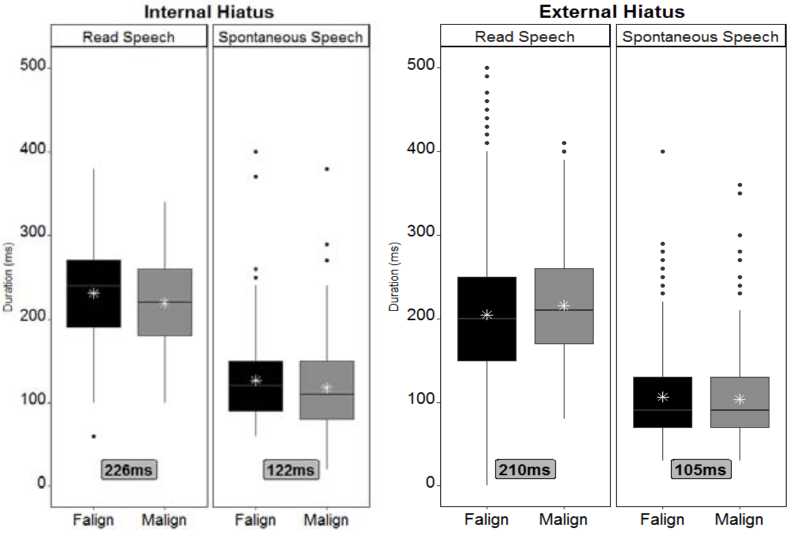

Figure 1. Duration patterns of IntH and ExtH according to speech style and data alignment.

Both internal and external hiatus manifest a similar reduction in the temporal domain from RS to SS (104ms for IntH and 105ms for ExtH). Also, our data show that, independent of speech style, ExtH is shorter than IntH:

Table 1

Mean duration of IntH and ExtH in RS and SS

	RS	**SS**
IntH	226ms	122ms
	st.dev$_{Falign}$ = 52	st.dev$_{Falign}$ = 52
	st.dev$_{Malign}$ = 54	st.dev$_{Malign}$ = 53
ExtH	210ms	105ms
	st.dev$_{Falign}$ = 82	st.dev$_{Falign}$ = 53
	st.dev$_{Malign}$ = 59	st.dev$_{Malign}$ = 53

The temporal differences between IntH and ExtH in the two speaking styles were analysed according to stress position, distribution in the word, number of syllables, vowel height, and place of articulation (variables added manually to the forced aligned output). In this article we draw attention on stress and syllable number.

We looked only at primary stress, and considered monosyllabic content words as stressed and monosyllabic function words as unstressed, unless placed in a prominent position.

For IntH there are four contexts:

- *the first vowel is stressed* (/fotogra'fie/ ‘photography’, n = 71);
- *the second vowel is stressed* (/lu'əm/ ‘we take’, n = 164);
- *hiatus in pretonic position* (/greu'tate/ ‘weight’, n = 105);
- *hiatus in post-tonic position* (/dele'gaʦie/ ‘delegation’, n = 80).

As for ExtH, out of the four possible outcomes (first or second vowel stressed, both vowels (un)stressed), only three surfaced in our data:

- *the first vowel is stressed* (/ʧe'va organi'zat/ ‘something organized’, n = 45);
- *the second vowel is stressed* (/'data 'asta/ ‘this time’, n = 177);
- *the vocalic pair is unstressed* (/'unde ur'me̯azə/ ‘where it follows’, n = 602).

IntH was found in words ranging between two to seven syllables (/strəin^j/ ‘strangers’, n = 45; /roʃie/ ‘red’, n = 119; /deʧizia/ ‘the decision’, n = 130; /informaʦie/ ‘information’, n = 96; /imaʤinaʦie/ ‘imagination’, n = 7; /antikariatele/ ‘the antique stores’, n = 3).

In the case of ExtH, four groups were outlined:

- monosyllabic – monosyllabic (MM; /nu am/ ‘I don’t have’, n = 90);
- monosyllabic – plurisyllabic (MP; /de unde/ ‘from where’, n = 266);
- plurisyllabic – monosyllabic (PM; /pentru o/ ‘for a(n)’, n = 186);
- and plurisyllabic – plurisyllabic (PP; /vara asta/ ‘this summer’, n = 282).

For the last category another subdivision was made based on the number of syllables of the first word in the pair (four contexts, namely *2 – n*, *3 – n*, *4 – n*, and *5 – n* syllables).

We expect to find a difference in duration based on stress distribution, and shorter duration with increasing number of syllables, both of which are conditioned by speech context.

3.1. Duration of IntH and ExtH According to Primary Stress Distribution

The results from duration and primary stress distribution show a systematic reduction of IntH with stress on the first vowel, followed by hiatus with the second vowel stress, then pretonic and post-tonic positions were the shortest, both in RS and SS (confirming Aguilar's (1999) claim that IntH are longer when the first vowel is stressed, followed by hiatus with the second vowel under stress):

Table 2

Mean duration of IntH in RS and SS according to primary stress distribution

	RS	**SS**
IntH V1	268ms	152ms
	$st.dev_{Falign}$ = 30	$st.dev_{Falign}$ = 47
	$st.dev_{Malign}$ = 32	$st.dev_{Malign}$ = 52
IntH V2	254ms	129ms
	$st.dev_{Falign}$ = 36	$st.dev_{Falign}$ = 40
	$st.dev_{Malign}$ = 34	$st.dev_{Malign}$ = 44
IntH PrT	203ms	114ms
	$st.dev_{Falign}$ = 50	$st.dev_{Falign}$ = 44
	$st.dev_{Malign}$ = 44	$st.dev_{Malign}$ = 40
IntH PsT	164ms	93ms
	$st.dev_{Falign}$ = 27	$st.dev_{Falign}$ = 74
	$st.dev_{Malign}$ = 31	$st.dev_{Malign}$ = 71

The same pattern holds true for ExtH. Vowel sequences with the first vowel stressed are longer than vowel sequences with the second vowel stressed, all of which are longer than their unstressed counterpart:

Table 3

Mean duration of ExtH in RS and SS according to primary stress distribution

	RS	**SS**
ExtH V1	258ms	142ms
	$st.dev_{Falign}$ = 104	$st.dev_{Falign}$ = 31
	$st.dev_{Malign}$ = 65	$st.dev_{Malign}$ = 21
ExtH V2	222ms	128ms
	$st.dev_{Falign}$ = 59	$st.dev_{Falign}$ = 61
	$st.dev_{Malign}$ = 46	$st.dev_{Malign}$ = 58
ExtH unstressed	203ms	96ms
	$st.dev_{Falign}$ = 84	$st.dev_{Falign}$ = 49
	$st.dev_{Malign}$ = 59	$st.dev_{Malign}$ = 49

Due to unequal variances (p <.05 for the Levene test) and unequal sample size, a Brown-Forsythe and Welch test were conducted, with 'Duration' as the dependent variable, while 'Speech' (two levels – 'RS', 'SS') and 'Stress Position' (four levels for IntH – 'V1', 'V2', 'PrT', 'PsT'; three levels for ExtH – 'V1', 'V2', 'V1V2unstressed') were the two independent variables. Both tests revealed significant differences (p<.0001). A Games-Howell post-hoc test was performed which showed an interaction between the factors only in RS.

In short, we view these results as supporting our hypothesis: *duration patterns of IntH and ExtH vary according to primary stress distribution*. We observed that hiatus is longer in stressed contexts than in unstressed contexts.

The two hierarchies that emerge from our data are:

- V1 stressed > V2 stressed > PrT position > PsT position for IntH (*Figure 2*),
- V1 stressed > V2 stressed > V1V2 unstressed for ExtH (*Figure 3*).

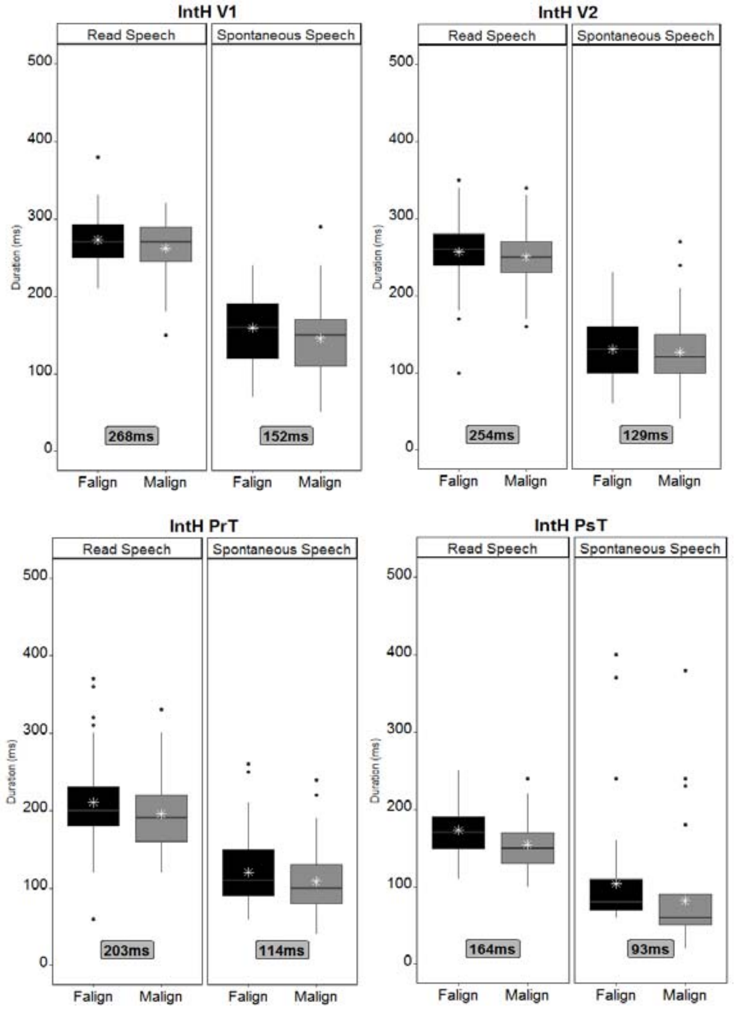

Figure 2. Duration patterns of IntH according to primary stress distribution.

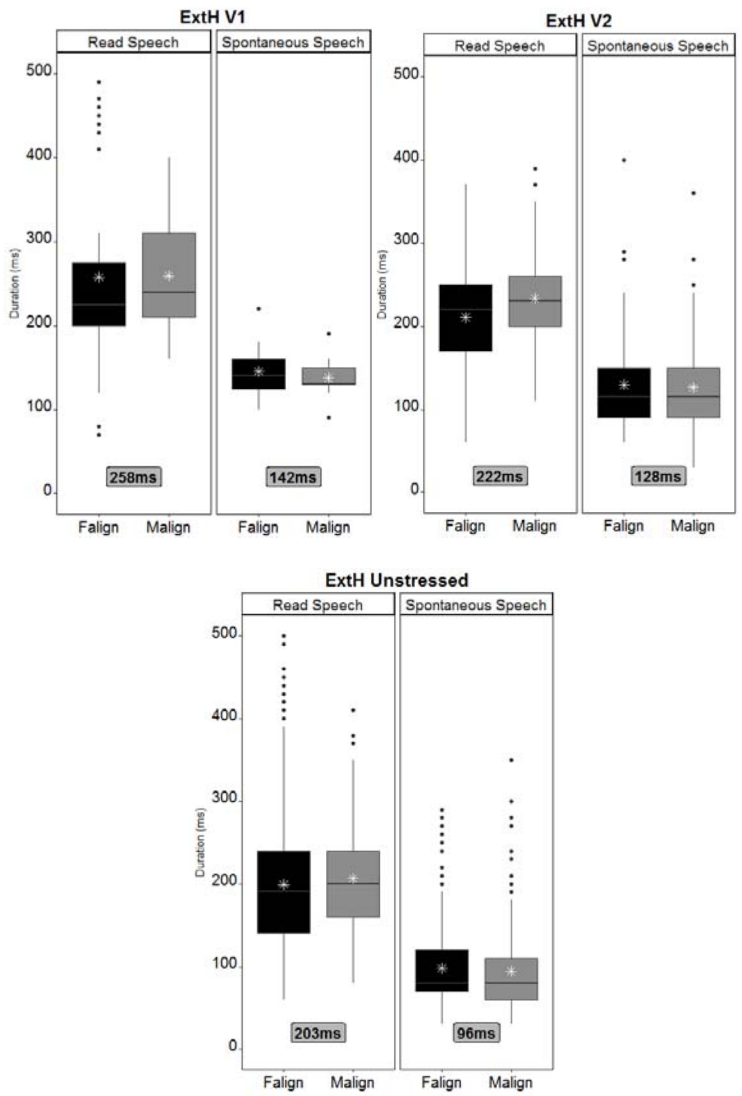

Figure 3. Duration patterns of IntH according to primary stress distribution.

3.2. Duration of IntH and ExtH According to Number of Syllables

Out of the four configurations of ExtH, MM and PM have the longest duration in RS, followed by MP and PP. This hierarchy is inversed in SS.

In the last configuration of PP (*5 – n* syllables), only three pairs were identified in the aligned corpus, namely 5 – 2 (/rədəʧinile əle̯a/ 'those roots', n = 9), 5 – 4 (/ɨntɨmplərile amuzante/ ‘amusing events’, n = 3) and 5 – 5 syllable words (/evenimente organizate/ ‘organized events’, n = 6). These outputs are underrepresented in SS. Also, the mean duration of 5 – 2 syllable words (96ms) is close to that identified in 5 – 5 examples (95ms), meaning that the duration of hiatus does not shorten gradually in SS (based on our corpora), thus leading to a week negative correlation.

In this context, a Pearson correlation test was carried out with the following results:

Table 4

Mean duration of IntH and ExtH in RS and SS according to syllable number

	RS	**SS**
IntH	-.47	-.94
ExtH_MP	-.76	-.50
ExtH_PM	-.76	-.92
ExtH_PP $\overline{2,n}$ syllables	-.86	-.70
ExtH_PP $\overline{3,n}$ syllables	-.62	-.86
ExtH_PP $\overline{4,n}$ syllables	-.60	-.80
ExtH_PP $\overline{5,n}$ syllables	-.88	-.21

IntH shortens more in SS (r = –.94) while ExtH shortens most in the PM configuration (r =–.92).

In sum, except for ExtH with 5 syllables in the first word (data were unbalanced), there is a strong negative correlation confirming our hypothesis that *hiatus shortens as the number of syllables increases*.

4. CONCLUSIONS AND FURTHER STUDIES

In this paper we compared IntH and ExtH in two speaking styles (RS and SS). Data were collected from one speaker (140 min) based on manual and forced alignments.

We conclude that duration patterns of VV-sequences vary according to primary stress distribution and number of syllables both of which are conditioned by speaking style. This analysis opens up discussion in relation to modelling gradient phonetic and phonological distinctions between IntH and ExtH. By comparing duration patterns of IntH an ExtH depending on context of speech (RS vs. SS) we can better understand a typology of hiatus resolution in Romance languages.

Future research topics will focus on extending the corpus by including speakers from other Romanian dialects and correlating duration patterns of IntH and ExtH with formant trajectories.

REFERENCES

Alba, M. C., 2006, “Accounting for variability in the production of Spanish vowel sequences”, in: N.Sagarra, A. J. Toribio (eds), *Selected proceedings of the 9th Hispanic Linguistics Symposium*, Somerville, MA:Cascadilla Proceedings Project, 273–285.

Adda-Decker, M., 2006, “De la reconnaissance automatique de la parole a l’analyse linguistique des corpus oraux”, *Actes des XXVI Journées d’étude sur la Parole* (JEP 2006), 12–16 June, Dinard, 389–400.

Aguilar, L., 1999, “Hiatus and diphthong: Acoustic cues and speech situation differences”, *Speech Communication*, 28, 57–74.

Baltazani, M., 2006, “Focusing, Prosodic Phrasing and hiatus resolution in Greek”, *Laboratory Phonology*, 8, 473–494.

Boersma, P., D. Weenink, 2011, *Praat: doing phonetics by computer* (version 6.0.37) [Computer program].

Cabré, T., P. Prieto, 2006, “Exceptional Hiatuses in Catalan and Spanish”, *Optimality-Theoretic Studies in Spanish Phonology*, 99, 205–238.

Casali, R. F., 1996, *Resolving Hiatus*, ms., PhD thesis, UCLA.

Chițoran, I., 2002a, *The Phonology of Romanian: A Constraint-Based Approach*, Berlin & New York: Mouton de Gruyter.

Chițoran, I., 2002b, “A Perception-Production Study of Romanian Diphthongs and Glide-Vowel Sequences’, *Journal of the International Phonetic Association*, 32/2, 203–222.

Chițoran, I., J. I. Hualde, 2007, “From hiatus to diphthong: the evolution of vowel sequences in Romance”, *Phonology*, 24, 37–75.

Colantoni, L., A. Limanni, 2008, “Where are hiatuses left? A comparative study of vocalic sequences in Argentine Spanish”, paper presented at the 38th Linguistic Symposium on Romance Languages, Urbana-Champaign, IL.

Ernestus, M., N. Warner, 2011, “An introduction to reduced pronunciation variants”, *Journal of Phonetics*, 39, 253–260.

Hualde, J.I., M. Prieto, 2002, “On the diphthong/hiatus contrast in Spanish: some experimental results”, *Linguistics*, 40/2, 217–234.

Kabak, B, 2007, “Hiatus Resolution in Turkish: an underspecification account”, *Lingua*, 117, 1378–1411.

Kainada, E, 2007, “Prosodic Boundary Effects on Duration and Vowel Hiatus in Modern Greek”, ICPhS, XVI, 1225–1228.

Niculescu, O., 2015, “Hiatul – Delimitări teoretice şi terminologice” [Hiatus – a theoretical and terminological account], *Studii şi cercetări lingvistice*, LXVI, 237–245.

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

Niculescu, Oana, 2018b, “Designing and exploring a speech corpus”, in: M. Candea, M. Causa, A. Ciugureanu, H. Quanquin, M. Vlad (eds), *International Journal of Cross-Cultural Studies and Environmental Communication*, 7, 1, Constanța, Editura Universitară şi Asociația pentru Dezvoltare Interculturală (ADI), 121–126.

Ohala, J. J., 1996, “The connection between sound change and connected speech processes”, *Arbeitsberichte* (AIPUK), University of Kiel, 201–206.

Prince, A. S, P. Smolensky, 1993, *Optimality Theory: Constraint Interaction in Generative Grammar*, ms., Rutgers University and the University of Colorado, Boulder.

R Core Team, 2013, *R: A language and environment for statistical computing. R Foundation for Statistical Computing*, Vienna, Austria, URL http://www.R-project.org/.

St-Amand, A. B., 2012, *Hiatus and hiatus resolution in Québécois French*, ms., PhD thesis, University of Toronto.

Vasilescu, I., B. Vieru, L. Lamel, 2014, “Exploring pronunciation variants for Romanian speech-to-text transcription‘, *Proceedings of SLTU-2014*, 161–168.