Connected speech in Romanian: Exploring sound change through an ASR system

Ioana Chitoran[a], Ioana Vasilescu[b], Bianca Vieru[c], Lori Lamel[b]

[a] Université Paris Diderot and Laboratoire de Phonétique et Phonologie (UMR 7018)
[b] LIMSI CNRS [c] Vocapia Research

Abstract

This study explores the hypothesis that traces of sound change can be found in connected speech processes as synchronic variation. To test this hypothesis we study the variation in connected speech through the output of an ASR system (Adda-Decker and Lamel, 1999). The case study is the vowel alternation [e]–[ʌ] after labials, encountered in Romanian historically, as part of a relatively sporadic sound change with unclear conditioning, as well as synchronically, as a phenomenon commonly observed in continuous speech. The specific question we address is whether the synchronic phenomenon in Romanian continuous speech is simply the result of general vowel reduction, or may reflect the sound change, with its apparent sensitivity to a labial consonantal context. We focus on the prepositions [pe] 'on' and [de] 'of', chosen specifically for their status as function words, more prone to vowel reduction than lexical words. We compare the ASR system's choice of two transcription variants and their respective contexts. The results of this comparison suggest that the synchronic variation is grounded in the historical sound change, thus motivating a future systematic study of controlled speech, to determine the exact role of segmental contextual factors. The main result is that the selection of transcription variants is significantly different for the two prepositions. This allows us to better understand the conditioning environment of the relevant sound change, and reveals the language-specific articulatory and co-articulatory settings and dynamic patterns that may have led to it, as they emerge and can be observed in continuous speech.

1 Introduction

The goal of this study is to explore the relationship between synchronic variation and historical sound change as it is specifically manifested in connected speech processes. The relationship between synchronic variation and sound change is unquestionable, as shown by Ohala (1989; 1996), Harrington (2012), Solé and Recasens (2012), among others. Two main approaches have been used. In some of this work, common sound changes have been studied by being replicated in the laboratory, under controlled conditions. In other type of work, inter- and intra-speaker variation have been studied in large speech corpora. Both approaches are valuable. Large scale data are particularly appropriate when the degree of variation that is targeted is only found in naturally occurring continuous speech, and cannot be elicited in a laboratory setting. In this study we are focusing on what we can learn about the evolution of a sound system from connected synchronic speech. To what extent can traces of sound change be found in connected speech? Our main hypothesis relies on the assumption that coarticulation patterns are language-specific (e.g., Manuel 1999; Beddor et al., 2002). If that is the case, it follows that connected-speech processes can also have language-specific manifestations.

We hypothesize that traces of sound change can be found in connected speech processes, and that language-specific coarticulatory settings and dynamic patterns can emerge and can be observed in continuous speech. In order to test our hypothesis we propose to analyze the variation occurring synchronically in a collection of Romanian continuous speech corpora. The novelty of our approach lies in resorting to automatic data alignment by a speech recognition system developed for Romanian (Vasilescu et al., 2014). Besides the practical benefits of forced alignment, we use the ASR system itself as a linguistic analyzer (Adda-Decker and Lamel, 1999). We examine the ASR system's choice between two transcription variants, namely /e/ and /ʌ/ for lexical /e/ in the prepositions [pe] 'on' and [de] 'of', as well as in controls: word-internal sequences [pe] and [de] and [e] encountered in all contexts where it is not preceded by [p] or [d]. We describe the method and the speech recognition system in section 3. This approach has recently been tested on the same Romanian speech corpus to study the marginal contrast among the two central Romanian vowel phonemes /ɨ̇/ and /ʌ/ (Renwick et al., 2016).

2 The data: [e] – [ʌ] alternation in Romanian

The data we propose to study involves the interaction of two vowels, the mid front [e] and the mid central [ʌ]. Historically, an [e] – [ʌ] alternation is attested as a rather sporadic change of vowel backing from Latin to Romanian (Nandriș 1963, Vasiliu 1968). Synchronically, in fluent speech the mid front [e] of function words such as [pe] 'on', [de] 'of', or [se] reflexive clitic pronoun is realized as [ʌ]. The phonemic vowel system of Romanian is presented in Table 1 below.

[-back]	**[+back]**	
i	ɨ̇	u
e	ʌ	o
ea	a	oa

Table 1. The Romanian phonemic vowel system (Chitoran 2001)

In Romanian, the central vowels /ɨ̇/ and /ʌ/ have been argued to pattern phonologically as back vowels (Chitoran 2001). We will therefore refer to the sound change as vowel backing rather than reduction. The diphthongs /e̯a/ and /o̯a/ have dual status in the Romanian system: there is evidence for both underlying and derived diphthongs in the language. They have been argued to pattern as low front and low back vowels, respectively (Chitoran 2001).

2.1 Sound change [e] > [ʌ]

The vowel [ʌ] in Romanian has several historical sources. Some of its occurrences are attributed to unstressed vowel reduction (Vasiliu 1968), traced back to the transition from Balkan Romance to Common Romanian. Unstressed (unrounded) vowels, whether part of the root or of the desinence, evolve into [ʌ], as shown in example (1):

(1) Unstressed vowel reduction:

LATIN >	ROMANIAN	
CASA(M)	ˈkasʌ	'house'
PARENTE(M)	pʌˈrinte	'parent'
SANITATE(M)	sʌnʌˈtate	'health'

The historical sources of [ʌ] that are of particular interest to us are instances of vowel backing, described below in (2) and (3). Vowel backing occurs regardless of stress and involves two consonantal environments. One source of [ʌ] is attested in the context of the Latin trill in word-initial position. In this environment, vowel backing ([i] > [ɨ]; [e] > [ʌ]) is attributed to the predorsum lowering and postdorsum retraction involved in the production of trills (Recasens 2014). In Romanian, [r] affects vowels both progressively and regressively, as illustrated in (2) and (3), respectively (Vasiliu 1968; Alkire and Rosen 2010).

(2) Progressive vowel backing ([i] > [ɨ]; [e] > [ʌ]) in trill context:

RĪPA	ˈrɨpʌ	‘river bank’
RĪDĬT	ˈrɨde	‘s/he laughs’
RĪMA	ˈrɨmʌ	‘worm’
RĪVU	ˈrɨw	‘river’
RĔUS	ˈrʌw	‘bad’

(3) Regressive vowel backing (to [ʌ]) in trill context:

	Indefinite	Definite	
PĬRU(M)	ˈpʌr	ˈpʌr-u-l	‘pear-tree’
PĬLU(M)	ˈpʌr	ˈpʌr-u-l	‘hair’
VĒRUS	ˈvʌr	ˈvʌr-u-l	‘cousin’
MĒLU(M)	ˈmʌr	ˈmʌr-u-l	‘apple / apple-tree’

In (3), vowel backing cannot be attributed exclusively to [r], because other regularities can be observed. For example, the masculine desinence vowel [-u] surfaces in the definite forms, where it precedes the definite marker [-l]. As we see by comparing (3) and (4), the vocalic context of the syllable added by an inflectional marker becomes relevant in synchronic alternations. The examples in (4) show back vowels in the singular roots, before the (underlying) back [-u] desinence, alternating with the front vowel [e] in singular forms before a front plural marker [-j] or [-e]. These types of alternations have been analyzed as metaphony patterns (Chitoran 2001; Marin 2007; Renwick 2012) because they affect the stressed vowel of the root. They are pervasive in Romanian morpho-phonology.

(4) Synchronic metaphonic alternations:

Singular	Plural	
ˈpʌr	ˈper-j	‘pear-tree’
ˈvʌr	ˈver-j	‘cousin’
ˈmʌr	ˈmer-j	‘apple-tree’
ˈmʌr	ˈmer-e	‘apple’ (fruit)

Finally, a third generalization of the sound change [e] > [ʌ] refers to the environment of a back vowel in the following syllable, and an immediately preceding labial. According to Vasiliu (1968), [e] becomes [ʌ] after a labial and before a syllable containing a back vowel. Several examples are given in (5).

(5) Vowel backing (to [ʌ]) in labial context:

FĒTU(M)	ˈfʌt	‘boy’
VĬDĔŌ	ˈvʌd	‘I see’
VETERANUS	bʌˈtrɨn	‘old’
METŬLA	ˈmʌturʌ	‘broom’
PECCATU(M)	pʌˈkat	‘sin’
PAVIMENTU	pʌˈmɨnt	‘earth’
VERSO	ˈvʌrs	‘I pour’
HOSPĬTIU	osˈpʌts	‘feast’
MEDULLA	ˈmʌduvʌ	‘marrow’
PEDUC(U)LU	pʌˈduke	‘lice’
FEBRUARIU	fʌuˈrar	‘February’

When we consider the labial context in the examples in (3) and (5) we see that the labial seems to have a lowering effect on the vowel. Unlike in (2), the expected high vowel [ɨ] never surfaces for a Latin high front vowel; [ʌ] surfaces instead.

A broader view of the sound changes from Latin to Romanian reveals an interaction between vowel backing and the diphthongization/vowel lowering that occurs under stress, and through which the diphthongs [e̯a] and [o̯a] developed in Romanian. As mentioned earlier, these diphthongs have been argued to pattern as low vowels (see Table 1). In (6) we consider some examples where vowel backing [e] > [ʌ] interacts with vowel lowering under stress [e] > [e̯a].

(6) [e] > [e̯a] > [a]

FĒTA (Lat.)	ˈfe̯atʌ > ˈfatʌ	‘girl’
MENSA (Lat.)	ˈme̯asʌ > ˈmasʌ	‘table’
nevěsta (Sl.)	neˈve̯astʌ > neˈvastʌ	‘wife’

According to the Latin and Slavic examples above, vowel lowering to [e̯a] had to precede vowel backing. For this reason, the vowel surfacing in the synchronic forms is the low [a] instead of the expected mid [ʌ]. Note that such examples also involve a preceding labial.

Vowel backing in a labial context is attested only in the transition from Common Romanian to Daco-Romanian dialects, but does not occur in the other main dialects – Aromanian, Megleno-Romanian, Istro-Romanian. The cross-dialectal comparison is illustrated by the examples in (7), from Vasiliu (1968).

(7) Other dialects maintain [e] and [e̯a] after labials:

Aromanian; Megleno-Romanian [e, e̯a]		Istro-Romanian [e]	
mer	‘apple-tree’	fet	‘boy’
per	‘pear-tree’	meturʌ	‘broom’
ved	‘I see’	pekɔt	‘sin’
fe̯atʌ	‘girl’		
me̯asʌ	‘table’		
nive̯astʌ	‘wife’		

Progressive vowel backing after a trill does, however, affect all dialects. Relevant examples are presented in (8), based on Nandriș (1963).

(8) All dialects undergo vowel backing after a trill:

	Daco-Romanian [ʌ,ɨ]	Megleno-Romanian [ɔ]	Aromanian; Istro-Romanian [ʌ,ɨ]	
RĔUS	rʌw	rɔw	arʌw	'bad'
RĪVU	rɨw	rɔw	arɨw	'river'
RĪPA	rɨpʌ	rɔpʌ	arɨpʌ	'river bank'

Another interesting historical aspect of the labial environment is that vowel backing in this context seems to have reversed direction over time. In 16th century texts, many more lexical items are encountered with the orthographic symbol <ă> used for [ʌ], than are attested in modern Romanian. Densușianu (1901) gives the following examples: <avăm> for Modern Standard Romanian <avem> [avem] 'we have', <iubăscu> for <iubesc> [jubesk] 'I love', <trimăs> for <trimes> [trimes] 'sent'.

To summarize the historical facts, it appears that the forms which have reliably preserved the back vowel in modern Romanian are the ones where all three environments are present: (i) a preceding labial consonant, (ii) a trill, and (iii) a back vowel (e.g., [mʌr-u-l] 'apple-def'). Summarizing the effects of each environment, we understand that the trill (or fortis [r] in Romanian) favors vowel backing through the post-dorsum retraction it involves. We also understand the backing effect as anticipatory coarticulation with a back vowel in the following syllable. What is less clear is the role of the labial in relation to backing. It is known that lip rounding, by lengthening the vocal tract, lowers all formants. Commonly, formant lowering by a labial consonant affects a following vowel, resulting in a labial vowel. If lip rounding were an active trigger of coarticulation in the cases examined here, we would expect vowels in these contexts to also become rounded, rather than just back.

To pursue this question we turn to the variation occurring in continuous speech in standard Romanian, which involves the same vowels, [e] and [ʌ].

2.2 Variation in continuous speech involving [e] and [ʌ]

The specific variation we focus on is encountered in function words, specifically the prepositions [de] 'of/from' and [pe] 'on'. A few representative examples are shown in (9).

(9) Variation in continuous speech

pe ~ pʌ ˈuʃʌ	'on the door'
pe ~ pʌ ˈtine	'you' sg. Accusative
de ~ dʌ ˈunde	'from where'
de ~ dʌ ˈmine	'by/of me'

We are not concerned here with the sociolinguistic aspects of this variation, nor with regional factors that characterize it. While both of these issues are interesting, in the present study we focus on the conditioning factors of the variation involving the same vocalic segments as the sound change. The variation illustrated in (9) is sufficiently salient and broadly distributed to both standard and non-standard varieties of Romanian to warrant its close analysis, and to justify drawing general conclusions from the results obtained.

We consider two hypotheses regarding variation in continuous speech. First, the variation encountered in function words in continuous speech may be due to vowel reduction or vowel centralization which normally affects function words in fluent speech, in all languages where it has been studied (Torreira and Ernestus, 2010; Meunier and Espeser, 2011). In this case the variation should be observed across the board, and should not be sensitive to specific

contextual factors. Alternatively, the synchronic variation may echo the historical sound change. If it is due to the same coarticulatory tendencies as the historical vowel backing, it is expected to show some sensitivity to the contextual factors discussed in section 2.1. We expect to see, specifically, evidence for anticipatory coarticulation with a back vowel, and/or sensitivity to a preceding labial, and/or sensitivity to to [r] in its context.

3 Working with the ASR system

We are using a 7-hour corpus of broadcast speech acoustically and statistically explored through the automatic alignment produced by an ASR system. The data were gathered from radio and television shows, and includes data from 141 male and female adult speakers. The corpus includes read or (semi) prepared speech, as well as more spontaneous interactions from televised debates. The speech is representative of the standard variety of Romanian, based on the Southern dialect. Portions with a large amount of overlapping speech were, as much as possible, eliminated, as well as foreign or regional accents, and noisy backgrounds. The data were automatically aligned and segmented into words and phones using the system developed by Vasilescu et al., 2014 for transcribing a large corpus of continuous speech in Romanian. The system generates automatic alignment and segmentation of the data into words and phones. The ASR system is based on a set of 29 phones, which include 20 consonants, 2 glides, 7 vowels and a special symbol for silence, presented in Table 2.

IPA	Ex. Romanian	IPA	Ex. Romanian
p	**p**as	b	**b**an
t	**t**are	d	**d**ac
k	**c**al	g	**g**ol
m	**m**ic	n	**n**or
f	**f**oc	v	**v**al
s	**s**are	z	**z**id
h	**h**orn	ts	**ț**ară
r	**r**epede	l	**l**ung
ʃ	**ș**arpe	ʒ	**j**ar
ʧ	**c**er	ʤ	**g**er
a	**a**pa	e	**e**rou
i	**i**nsula	o	**o**ra
u	**u**dă	ə	ud**ă**
ɨ	**î**nspre		
o̯a	f**oa**rte	j	**i**apa
e̯a	m**ea**	w	da**u**
-	*silence*	-	*breath*
-	*filler*		

Table 2. Phone set used for the automatic transcription system

For the present study we are using an improved version of the system, with acoustic and language models trained as in Renwick et al. (2016). All acoustic models are built in a semi-supervised manner using approximately 370 hours of untranscribed audio (Gauvain, Lamel and Adda 2002; Lamel and Vieru 2010). In our study the system is allowed to align different variants for lexical /e/.

An experiment relying on pronunciation variants allows us to go beyond an acoustic analysis of production data, which would primarily document a general trend toward centralization due to shorter duration in continuous speech (Gendrot and Adda-Decker 2005),

especially in the case of function words, which is our focus. Instead, the use of pronunciation variants has the advantage of combining acoustic decoding and the analysis of contextual distribution (Renwick et al., 2016). In this work we adopt the methodology of pronunciation variants because it allows for a better estimation of the combined effects of acoustic patterns and context.

The proposed variants for lexical /e/ are /e/ or /ʌ/. The resulting transcriptions are then sorted according to context. We will count the occurrences of the two proposed variants and correlations between the selected variant and its context:

- % of /e/ in the prepositions [pe], [de] that was transcribed as /ʌ/
- % of /e/ that was transcribed as /ʌ/ in other contexts (i.e., not prepositions) preceded by /p/ and /d/
- % of /e/ that was aligned as /ʌ/ in all other contexts (not in prepositions, and not preceded by /p/ or /d/)

Adopting the approach proposed by Adda-Decker and Lamel (1999), we study the extent to which variation can be linked to the variants selected by the ASR system run on a long sample of continuous speech. It is hypothesized that an ASR system will encounter difficulties precisely in loci of high variation. This approach therefore allows us to identify the contexts where variation occurs, and to determine the broad acoustic manifestations of these contexts. Moreover, it allows us to assess whether the observed variation is strictly contextually predictable, or tends to be generalized. The advantage of this method lies in providing a natural setting for the study of variation that is characteristic of continuous speech. However, the conclusions that we can draw based on the analysis of the forced selection of variants by an ASR system can only be considered preliminary, due to the uncontrolled nature of the large speech database. Here we test preliminary linguistic hypotheses on naturally occurring speech, but in order to fully characterize the sound change we are interested in, and to determine its detailed properties, further analysis of fine phonetic properties in carefully controlled speech is needed.

3.1 Hypotheses

Three hypotheses are tested, referring to the preceding and the following contexts.

- H1: If what occurs in continuous speech is solely vowel reduction in function words, the prepositions [de] and [pe] are expected to be realized as [dʌ] and [pʌ] respectively, to a comparable extent, regardless of segmental context.
- H2: If what occurs in continuous speech is vowel reduction reinforced by back vowel harmony and/or by the presence of [r], the prepositions [de] and [pe] are expected to be realized as [dʌ] and [pʌ] respectively, to a comparable extent, when followed by a back vowel and/or a trill in the following syllable.
- H3: If what occurs in continuous speech is vowel reduction reinforced by back vowel harmony, [r] coarticulation, as well as labial coarticulation, then:
    - [pe] is expected to be realized as [pʌ] more often than [de] is realized as [dʌ]
    - and this effect may be reinforced before a back vowel or [r] in the next syllable.

3.2 Results

The first comparison that was made involved all the instances of *pe* and *de* extracted from the corpus, whether they occurred in prepositions or not, and without considering the following segmental context. This overall comparison revealed that the distribution of the two transcription variants, /e/ and /ʌ/, is not quite the same across *de* and *pe*: $X^2(1) = 3.99$, $p = .045$. This difference motivates a series of subsequent comparisons which we present here,

meant to identify possible sources of this asymmetry in the selection of the transcription variants.

We next considered the two prepositions, separately from other occurrences of the sequences [de] and [pe]. In our corpus, the preposition count is unbalanced. The corpus contains many more occurrences of the preposition [de] (1287), than of the preposition [pe] (319). We found that the distribution of the two transcription variants (Table 3) is quite different across the two prepositions: $X^2(1) = 6.18$, $p = .012$.

PREPOSITIONS Variants	pe	de
/e/	58.4%	75.9%
/ʌ/	41.5%	24%

Table 3. Transcription counts of the prepositions [de] and [pe] with two allowed variants, /e/ and /ʌ/

The percentage values in Table 3 show that, overall, the /ʌ/ transcription variant is selected more often for [pe] (41.5%) than for [de] (24%).

We consider next the contextual role of [r] in this distribution. We see that the preference for the back vowel variant in [pe] is maintained in the presence of [r] in the following context (Table 4): the /ʌ/ variant is still selected significantly more often for the preposition [pe] (40%) than for [de] (21.5%): $X^2(1) = 7.21$, $p = .007$.

PREPOSITIONS Variants before /r/	pe	de
/e/	59.9%	78.5%
/ʌ/	40%	21.5%

Table 4. Transcription counts of the prepositions [de] and [pe] with two allowed transcription variants, /e/ and /ʌ/, as a function of /r/ in the following syllable

We can only consider here the context of a following [r], which is less relevant for a sound change. In Romanian, as shown in (2), vowel backing is attested as a carryover effect from a preceding [r], and less clearly as an anticipatory effect. In other Romance languages, as well, the favored coarticulatory direction for a trill is progressive, possibly due to the aerodynamics requirements imposed by its production (Recasens 2014). Moreover, in modern Romanian, [r] is more commonly realized as a tap rather than a trill, especially in a prosodically weak position. Taps involve less postdorsum retraction than trills, therefore the coarticulatory backing effect may not be as strong.

Indeed, the results suggest that, regardless of the presence of [r] in the following syllable, the vowel in the preposition [pe] is more often transcribed as back than the same vowel in the preposition [de]. The preference for the /ʌ/ transcription variant is significantly higher in [pe], and this does not change when we consider only the context of a following /r/.

The second contextual factor, the quality of the following vowel – front or back – does not make a difference in the choice of the transcription variant. Regardless of whether the following vowel is front or back, there is a strong preference for the selection of the /ʌ/ variant in [pe] compared to [de] (Table 5).

PREPOSITIONS Aligned variants	pe	de
/e/ front V	72%	85.2%
/e/ back V	55.7%	70.3%
/ʌ/ front V	27.9%	14.7%

/ʌ/ back V	44.2%	29.6%

Table 5. Transcription counts of the prepositions [de] and [pe] with two allowed variants, /e/ and /ʌ/, as a function of the following vowel quality

Thus, even in a front vowel context, the /ʌ/ variant is more often selected for [pe] than for [de] ($X^2(1) = 4.44$, $p = .035$), and the same is true in a back vowel context ($X^2(1) = 3.97$, $p = .04$). Figure 1 shows that the /ʌ/ variant is selected more often for [pe] than for [de] regardless of the vowel context.

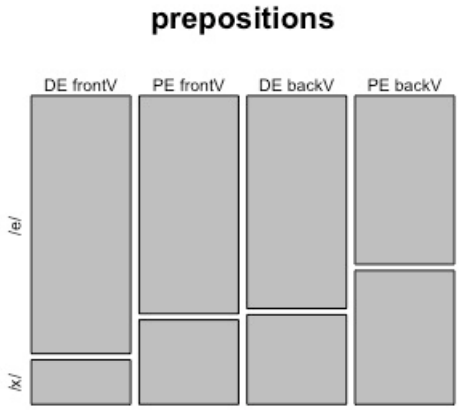

Figure 1. Distribution of the transcription variants /e/ and /ʌ/ across the prepositions [pe] and [de] in front vowel and back vowel contexts (/ʌ/ = /x/ in the figure)

This result suggests that the coarticulation with a following back vowel does not, by itself, predict the choice of the back transcription variant. The context of the preceding labial does make a difference. Even though the /ʌ/ variant is significantly more often selected in the context of a following back vowel both for [de] ($X^2(1) = 5.60$, $p = .001$), and for [pe] ($X^2(1) = 5.07$, $p = .02$), the choice of /ʌ/ is still highest for [pe]. This means that the labial itself is a good predictor of the choice of the /ʌ/ variant.

We next exclude the prepositions, and we consider the occurrences of [de] and [pe] as sequences elsewhere in the corpus. The data for sequences are more balanced. We counted 637 occurrences of the sequence [de] in the corpus, and 570 occurrences of the sequence [pe]. Overall, as for the prepositions, the distribution of the two transcription variants (Table 6) is not the same across the sequences [pe] and [de]: $X^2(1) = 4.23$, $p = .03$.

SEQUENCES Aligned variants	pe	de
/e/	65.4%	79.4%
/ʌ/	34.5%	20.5%

Table 6. Transcription counts of the sequences [de] and [pe] with two allowed variants, /e/ and /ʌ/

The transcription variant /ʌ/ is again selected more often when preceded by [p], but this time the preference is present only in a back vowel context. The result is highly statistically significant only before a back vowel in the following syllable ($X^2(1) = 13.97$, $p = .0001$), and not in the context before a front vowel ($p > .05$).

SEQUENCES Aligned variants	pe	de
/e/ front V	93%	98%

/e/ back V	28.6%	55.7%
/ʌ/ front V	6%	1.9%
/ʌ/ back V	71.3%	44.2%

Table 7. Transcription counts of the prepositions [de] and [pe] with two allowed variants, /e/ and /ʌ/, as a function of the following vowel context

Figure 2 shows that the /ʌ/ variant is more often selected when a back vowel follows in the next syllable. Moreover, the /ʌ/ variant is more often selected when the preceding consonant is a labial.

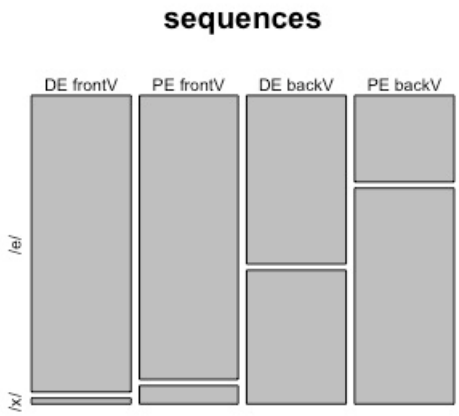

Figure 2. Distribution of the transcription variants /e/ and /ʌ/ across the sequences [pe] and [de] in front vowel and back vowel contexts (/ʌ/ = /x/ in the figure)

This result is clearly consistent with hypothesis 3.

To summarize, the results confirm the predicted sensitivity to two of the contexts considered – preceding labial, and following vowel quality. The coarticulatory effect of [r] is not clearly confirmed. In the case of prepositions, the effect of the preceding labial can be isolated from that of vowel quality. In both a back and a front vowel context the back vowel transcription variant was more often selected for the preposition [pe] than for [de]. In the case of sequences, the sensitivity is cumulative, in keeping with hypothesis 3. The back vowel transcription variant is more frequently selected for the sequence [pe], and this preference is reinforced in a back vowel context.

As already pointed out, these results can only be considered preliminary. They suggest that the variation encountered in the speech corpus is not just the manifestation of general reduction effects, and there may be a connection between the two realizations of the vowel as [e] or [ʌ], and the sound change involving the same segments. A more in-depth understanding of the sound change requires a careful acoustic study of balanced and controlled data. The results obtained so far serve primarily to justify pursuing such a study. They validate the importance of further testing the hypothesis on carefully controlled data.

We compared the values of the first two formants of the vowels in our corpus of continuous speech. Figure 3 shows F1 and F2 formant values plotted for [e] in the prepositions [pe] (in grey) and [de] (in black) for female speakers (left) and male speakers (right).

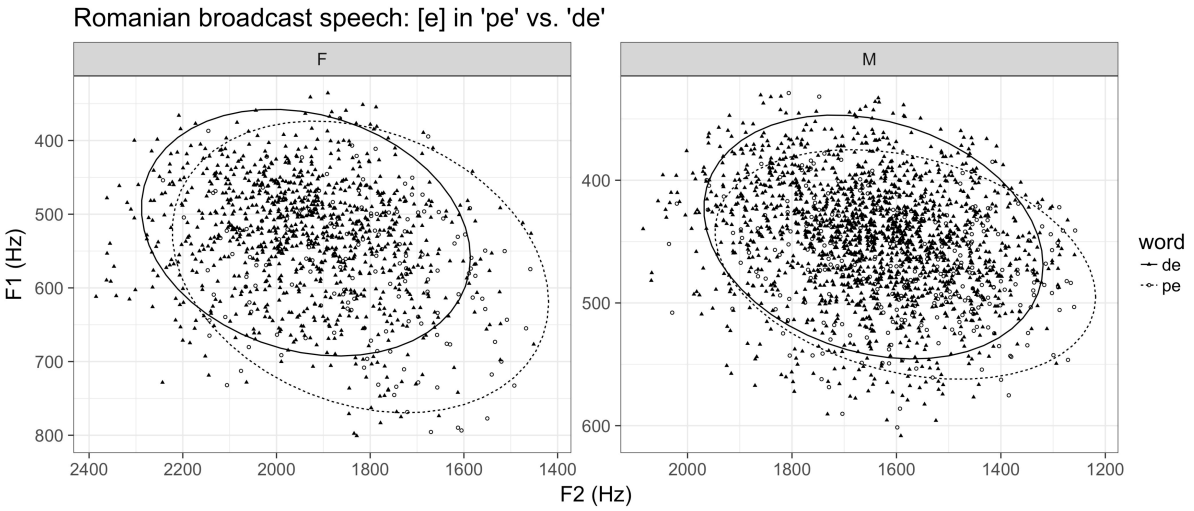

Figure 3. Vowel plots for [e] in the prepositions [pe] (grey) and [de] (black). Left: female speakers. Right: male speakers (Graph by Peggy Renwick).

F1 and F2 measures were taken at the vowel midpoint. Notice that for both male and female speakers, F2 values (y axis) for [pe] are lower than for [de], which indicates a less front vowel in [pe] than in [de]. The difference, evaluated in a T-test, is statistically significant. Welch two sample t-test revealed that F2 is higher in [de] (mean = 1937 Hz, SD=164) for female speakers [$t(201.3)=7.51$, $p < 0.0001$], as well as for male speakers (mean = 1648 Hz, SD=150), [$t(403.11=5.72$, $p < 0.0001$]. The data plotted in Figure 4 are raw data (in Hz), but similar results were obtained for normalized data, which were z-scored by speaker.

4 Discussion

To summarize, the /ʌ/ transcription variant was selected more often after /p/ than after /d/. In the case of the prepositions [pe] and [de], this preference was observed regardless of the following vowel context. The back variant /ʌ/ was predominantly selected for [pe] both before a back vowel in the following syllable and before a front vowel. The choice of the variant by the ASR system is consistent with the acoustic difference observed in the data. F2 of /e/ in [pe] is significantly lower than in [de] for both male and female speakers.

The back variant /ʌ/ was predominantly selected for [pe] before [r] in the following syllable, but it was also preferred in all the occurrences of the preposition, all contexts combined. In the case of the sequences [pe] and [de] occurring outside of prepositions, the back variant /ʌ/ was more often selected in a back vowel context, and in this context it was predominantly selected in [pe].

Because [pe] and [de] are treated differently by the ASR system, none of the results support hypotheses 1 and 2. The analysis of synchronic variation occurring in continuous speech suggests instead that the presence of a labial consonant favors backing of [e] to [ʌ]. In prepositions, the effect of the labial as manifested in the choice of the back transcription variant emerges independently of the following vowel quality. In the case of [pe] and [de] sequences, however, hypothesis 3 is confirmed: the effect of the labial is reinforced in the back vowel context (the [r] context was not examined for sequences).

Let us consider in more detail the coarticulatory effect of a labial on a vowel. It is known that labials show relatively low F2 frequency due to the lip closing gesture (Fant 1960; Recasens 2014). Recasens et al. (1997) found that carryover effects are more prominent in F2 frequency from [p] to high or mid vowels, rather than to [a]. This means that labials are expected to favor mid and high back rounded vowels in their vicinity. The Romanian examples in (3), (5) and (6), are only partly consistent with this prediction, as they

predominantly involve the vowel [e] being backed to [ʌ]. What is unusual about the Romanian data, however, is that the vowels do not undergo rounding, which usually accompanies backing. Indeed, in Recasens's (2014) comprehensive study of common types of sound change across varieties of Romance, the most common processes reported, triggered by labial consonants and affecting vowels, involve: progressive raising of [o] > [u], and regressive rounding assimilation [e,a,ə] > [o,u] or dissimilation [o,u] > [e,a,ə]. This detail motivates caution in asserting the role of the labial alone in this process.

The difference observed in the results between prepositions and sequences highlights the limit of the conclusions we can draw from this first study. The result for prepositions can be argued to be the more reliable one, because prepositions provide a more controlled condition, with fewer factors that can potentially affect the quality of the vowel. Most importantly, stress is consistently absent in prepositions, thus favoring vowel reduction. While the prepositions are always unstressed, the sequences [pe] and [de] include a mix of stressed and unstressed tokens. ASR transcriptions have not yet been analyzed separately with respect to the presence or absence of stress. This difference is very important, and may explain why the effect of the preceding labial is more evident in prepositions. For the more variable sequences, the effect of the labial only emerges within a specific coarticulatory context, that of a following back vowel. It is important, therefore, to consider only the unstressed [pe] and [de] sequences in the corpus, in order to compare their transcription to that of prepositions. We predict that the choice of the transcription variant will be more similar to the distribution observed for prepositions.

At the same time, one other difference between prepositions and sequences in our corpus gives more weight to the results obtained for sequences. Our corpus contains an unbalanced count of prepositions, with many more instances of [de] than of [pe], compared to the better balanced counts of [de] and [pe] sequences. It may be that, if the number of prepositions is balanced, the effect of the following vowel will become apparent for prepositions, as well. This is one of the tradeoffs encountered in working with large corpora: they provide large amounts of data containing speech phenomena that cannot be naturally elicited in the laboratory. At the same time, the data cannot be fully controlled, thus limiting the types of questions that can be asked. To address this limitation, we are currently planning an acoustic analysis of controlled speech which was elicited through guided conversations with pre-selected lexical items. The results of this study are needed to ultimately determine whether the labial is itself responsible for vowel backing, or whether anticipatory vowel-to-vowel coarticulation is the main predictor of vowel backing in [pe]. In the latter case, F2 of the vowel is predicted to lower more, under the combined coarticulation with the preceding labial and with the following vowel. Some facts consistent with this interpretation can be observed. The lexical items that synchronically show an alternation between [e] and [ʌ] always involve a preceding labial, as well as a back vowel in the following syllable. Consider, for example, [osˈpʌts-u-l] 'feast-def.' vs. [ospeˈts-ie] 'hospitality'.

Independently of these facts, Romanian vowels show a clear coarticulatory predisposition manifested in metaphony effects (Renwick 2012). Metaphony effects are more robust and regular than the vowel backing examined in the present study, but both types of phenomena attest to strong coarticulatory tendencies in the vocalic system, both historically and synchronically.

5 Conclusions and further directions

We have shown that the study of synchronic variation occurring in continuous speech can reveal traces of sound change. The automatic treatment of large speech corpora through forced alignment of transcription variants in an ASR system promises to be an efficient way of testing hypotheses about sound change. In spite of certain limitations which need to be

carefully considered, this method is most efficient in allowing, through preliminary testing, to refine hypotheses for subsequent, controlled studies. Automatic transcription tools allow us to study variation in continuous speech, and support the understanding that variation in synchronic continuous speech reflects past sound changes or foreshadows future ones. The results we obtained have undoubtedly helped us to refine our hypotheses, which will next be tested on carefully controlled data, free of confounding factors. We can then compare, for example, the normalized formant values of the vowels in the broadcast corpus (figure 3) with normalized formant values of the same vowels in controlled speech. We now know that our hypothesis is worth pursuing in speech elicited in identical recording conditions, complete with information about speech style, (socio)linguistic and regional dialectal background of the speakers, all of which are well known to have acoustic effects. The most reliable conclusions can be drawn based on results from both types of speech corpora: large continuous speech data and carefully controlled data.

References

Adda-Decker, Martine and Lori Lamel. Forthcoming. «Discovering speech reductions across speaking styles and languages». In M. Adda-Decker and N. Nguyen, eds., *Rethinking Reduction*. Mouton de Gruyter.

Adda-Decker, Martine and Lori Lamel. 1999. «Pronunciation variants across system configuration, language and speaking style». *Speech Communication* 29(2-4):83-98.

Alkire, Ti and Carol Rosen. 2010. *Romance Languages: A Historical Introduction*. Cambridge University Press.

Beddor, Patrice S., J.D. Harnsberger, S. Lindemann. 2002. «Language-specific patterns of vowel-to-vowel coarticulation: acoustic structures and their perceptual correlates». *Journal of Phonetics* 30: 591-627.

Boersma, P. and D. Weenink. PRAAT: doing phonetics by computer. http://www.praat.org/. 2015

Chitoran, Ioana. 2001. *The phonology of Romanian: A constraint-based approach*. Mouton de Gruyter 2001.

Densuşianu, Ovid. 1901. *Histoire de la langue roumaine*. Paris, Ernest Leroux Editeur.

Fant, Gunnar. 1960. *Acoustic Theory of Speech Production*. The Hague: Mouton

Gendrot, Cédric and Martine Adda-Decker. 2005. Impact of duration on F1/F2 formant values of oral vowels: an automatic analysis of large broadcast news corpora in French and German. Proceedings of Interspeech 2005, Lisbon, Portugal, November 2005. 2453-2456

Harrington, Jonathan. 2012. «The relationship between synchronc variation and diachronic change». In Abigail C. Cohn, Cécile Fougeron, Marie Huffman, eds. *Handbook of Laboratory Phonology*. Oxford University Press: Oxford. 321-332.

Manuel, Sharon. 1999. «Cross-language studies: Relating language-particular coarticulation patterns to other language-particular facts». In William J. Hardcastle and Nigel Hewlett, eds. *Coarticulation: Theory, data, and techniques*, Cambridge, Cambridge University Press: 179-198.

Marin, Ștefania. 2007. Vowel to vowel coordination, diphthongs and Articulatory Phonology. Yale University PhD dissertation.

Meunier, Christine and Robert Espesser. 2011. «Vowel reduction in conversational speech in French: The role of lexical factors». *Journal of Phonetics* 39:3, 271-278.

Nandriş, Octave. 1963. *Phonétique historique du roumain*. Paris: Klincksieck

Ohala, John J. 1989. «Sound change is drawn from a pool of synchronic variation». In: Breivik, LE, Jahr, EH, eds., *Language change: Contributions to the study of its cause*s. Mouton de Gruyter. 173-198.

Ohala, John J. 1996. «The connection between sound change and connected speech processes». *Arbeitsberichte (AIPUK 31)* Universität Kiel. 201-206.

Recasens, Daniel. 2014. *Coarticulation and sound change in Romance*. Amsterdam, John Benjamins.

Recasens, Daniel, Maria Dolors Pallarès, Jordi Fontdevila. 1997. «A model of lingual coarticulation based on articulatory constraints». *Journal of the Acoustical Society of America* 102, 544-561.

Renwick, Margaret E.L. 2012. *Vowels of Romanian: Historical, Phonological and Phonetic Studies*. Cornell University PhD dissertation.

Renwick Margaret E.L. 2014. *Phe phonetics and phonology of contrast: The case of the Romanian vowel system*. Mouton de Gruyter.

Renwick, Margaret E.L., Ioana Vasilescu, Camille Dutrey, Lori Lamel, Bianca Vieru. 2016. «Marginal contrast among Romanian vowels : Evidence from ASR and functional load». In *Proceedings of Interspeech 2016*, 2433-2437.

Solé, Maria Josep and Daniel Recasens. 2012. *The initiation of sound change: Perception, production, and social factors*. Amsterdam: John Benjamins.

Torreira, Francisco and Mirjam Ernestus. 2010. «The Nijmegen Corpus of Casual Speech». *Speech Communication* 52, 201-212.

Vasilescu, Ioana, Bianca Vieru, Lori Lamel. 2014. «Exploring pronunciation variants for Romanian speech-to-text transcription». In *Proceedings of SLTU* 2014, 161-168.

Vasiliu, Emanuel. 1968. *Fonologia istorică a dialectelor daco-române*. Editura Academiei.