1-1-2011

Phoneme Type Frequency in Romanian

Margaret E. L. Renwick
Cornell University, mer56@cornell.edu

Phoneme Type Frequency in Romanian

Abstract

This paper presents research into the relative frequencies of phonemes in Romanian, focusing on the high central vowel /ɨ/ to demonstrate how type frequency reflects its former allophonic status as an allophone of /a/ and later /ə/; and that the low type frequency of /ɨ/ correlates with the vowel's minimal expansion beyond its original allophonic environments.

The historical facts show that [ɨ] was allophonic in native words of Latin origin, emerging mainly through pre-nasal raising. Borrowings from Slavic, however, cannot be explained through allophony alone, and later borrowings from Turkish show a correspondence between Turkish [ɨ] and Romanian /ɨ/, indicating the vowel was on the verge of phonemic status. Although /ɨ/ is synchronically contrastive in Romanian, its contrastiveness is marginal; few minimal pairs separate the once-allophonic /ɨ/ and /ə/.

A type-frequency analysis allows us to examine the functional load of phonemes in modern Romanian, for comparison with the historical picture. Among the vowels of Romanian, type frequency varies widely; the least-frequent vowels are /ɨ/ and /ə/. The low type frequency of /ɨ/ in particular follows from its origins as an allophone, and from the circumstances of its phonemicization: /ɨ/ was originally conditioned in stressed syllables preceding a nasal, and also by a preceding /r/ or following /rC/. This is precisely the variety of phonological conditioning that can be shown through this type-frequency analysis.

With regard to /ɨ/, I argue that in the vowel's type frequency, we see little more than the phonological footprint of the processes that brought /ɨ/ into Romanian: the role of /ɨ/ has not expanded much beyond its original allophonic role.

Phoneme Type Frequency in Romanian

Margaret E. L. Renwick[*]

1 Introduction

This paper presents research into the relative frequencies of phonemes in Romanian, focusing on the high central vowel /ɨ/ to demonstrate how type frequency reflects its historical status as an allophone. It deals with the claim that the low type frequency of /ɨ/ in Romanian correlates with the vowel's minimal expansion beyond its original allophonic environments. We find that the vowel's distribution with respect to surrounding segments reflects the footprint of the phonological processes that created it. 0 shows minimal sets in which /ɨ/ contrasts with other vowels, demonstrating its synchronic phonemic status.

mână	/mɨnə/	'hand' (n.)	mină	/minə/	'mine' (n.)
râpă	/rɨpə/	'precepice'	rupă	/rupə/	'breaks' (3sg. subj.)
pâr	/pɨr/	*interj.*	păr	/pər/	'hair'
			par	/par/	'pole'

Table 1: Minimal sets with /ɨ/

A frequency analysis is useful in this case for several reasons. First, it measures how the functional load of the vowel space is divided up across the lexicon. Secondly, by examining the co-occurrences of segment pairs, we may detect phonological conditioning: if two segments co-occur more often than predicted by their respective frequencies, phonological processes might be the cause. This is the case in Romanian. Finally, breaking data down this way allows a new angle for examining phonological processes and their effect on the lexicon.

2 Romanian /ɨ/

2.1 A Recent Phonemic Addition

Among the phonemic vowels of Romanian, which appear in 0, /ɨ/ is the most recent addition. The origin of /ɨ/ in Romanian is controversial: Did it split from /ə/ (Rosetti 1986; Coteanu 1981; Vasiliu 1968), or was it introduced by Slavic loans (Hall 1974; Petrovici 1958; Mallinson 1988)? Petrucci (1999) argues that /ɨ/ cannot have come from Slavic *y, due to a lack of direct correspondences between loanword vowels and Romanian /ɨ/. Renwick (2009), whose analysis is adopted here, argues that /ɨ/ began as an allophone, and grew to phonemic status with help from loanwords.

Phonemic Vowel Chart		
i	ɨ	u
e	ə	o
	a	

Orthography		
i	â / î	u
e	ă	o
	a	

Table 2: Synchronic vowel phonemes of Romanian

2.2 The origins of /ɨ/

Several authors (Rosetti 1986; Coteanu 1981; Vasiliu 1968) demonstrate that /ɨ/ first arose as an

allophone of /ə/ in Romanian native vocabulary; thus, a central question is how the two split into separate phonemes. In native words, we find /ə/ and then /ɨ/ under the following conditions:

(1) Basic condition for */ə/:

$$*/a/ > */ə/ \; / \; C \underset{\text{[-stress]}}{\underline{\qquad}}$$

(2) Raising of stressed pre-nasal */a/:

$$*/a/ > */ə/ > /ɨ/ \; / \; \underset{\text{[+stress]}}{\underline{\qquad}} \begin{cases} n(C)(V) \\ mC \end{cases}$$

(3) Raising/centralization before a nasal:

$$/V/ > */ə/ > /ɨ/ \; / \; \underset{\text{[-low]}}{\underline{\qquad}} \begin{cases} n(C)(V) \\ mC \end{cases}$$

(4) Conditioning of /ɨ/ near /r/:

$$/V/ > /ɨ/ \; / \begin{cases} \underline{\quad}rC \\ r\underline{\quad} \end{cases}$$

The conditions in (1–4) above explain the conditioning of /ɨ/ in native words, from Latin, as seen in 0. Generally, we find that /ɨ/ arose from native words in which /a/ appeared before a nasal in stressed position; alternatively, /ɨ/ was conditioned after /r/ or before /rC/ (cf. Schulte 2005). Non-low vowels also emerged as /ɨ/, under similar circumstances. The changes that took place in these words represent, for the most part, a stage of Romanian development during which [ɨ] was not a phoneme.

Condition	Latin	Romanian		
2	manet	mâne	/mɨne/	'remains' (vb)
	lana	lână	/lɨnə/	'wool'
	campu	câmp	/kɨmp/	'field'
	granum	grâu	/grɨw/	'wheat'
	quantus	cât	/kɨt/	'how much?'
3	in	în	/ɨn/	'in'
	longu	lângă	/lɨŋə/	'beside, near'
	imperator	împărat	/ɨmpərat/	'emperor'
4	ridere	râde	/rɨde/	'laughs' (vb)
	*carnaceus	cârnat	/kɨrnat/	'sausage'
	tardivus	târziu	/tɨrziw/	'late'

Table 3: /ɨ/ in native words

2.3 The appearance of /ɨ/ in loan words

As loans entered Romanian, first from the Slavic languages and later from Turkish, we find many instances of /ɨ/ that cannot be explained by phonological conditioning (Renwick 2009). This indicates a gradual expansion of the environments in which /ɨ/ was permitted, eventually leading to phonemic status, as attested in the 16th Century (Rosetti 1986; Coteanu 1981; Vasiliu 1968; Sala 1976). In 0 and 0, we see examples of loanwords of Slavic and Turkish origin.

We find that while most words of Slavic origin have a conditioning environment for /ɨ/, a subset of words do not. When words were borrowed from Slavic, Romanian tended to use /ɨ/ in place of nasalized vowels, inserting a post-vocalic nasal consonant to retain the nasal element. Additionally, Romanian used /ɨ/ in place of some Slavic jers, specifically those which eventually underwent liquid-jer metathesis. These jers are assumed to have been phonetically different from other jers in Common Slavic, which do not emerge as /ɨ/ in Romanian. However, in certain cases, phonological conditioning cannot account for the presence of /ɨ/ in Slavic loan words, and we find some cases in which /ɨ/ in Romanian corresponds to its Slavic equivalent, <y>. From this we may

infer that Romanian speakers were beginning to perceive and use /ɨ/ as a distinct vowel, and that the phonological environments available for it were expanding.

Turning to words borrowed from Turkish, we do find instances of /ɨ/ that could have been conditioned, by a nasal or /r/. However, in a majority of the Turkish etymons for Romanian words containing /ɨ/, we find no conditioning environment for the vowel; instead, /ɨ/ corresponds to Turkish /ɨ/, particularly under conditions of stress. This characteristic only applies in a few words from Old Slavic, and the general lack of correspondences of this type has been used by Petrucci (1999) to argue for the native origins of /ɨ/ in Romanian.

In words from Turkish, the correspondence of /ɨ/ to /ɨ/ is common. Since these words were borrowed *later* than the bulk of Slavic loans in Romanian, we can hypothesize that the difference in vowel treatments across the loan classes is due to a general increase in speakers' familiarity with the vowel in non-conditioned contexts. These facts indicate that while Romanian did not distinguish between the phonemes /ɨ/ and /ə/ when the words in 0 entered its lexicon, speakers of the language retained vowels faithful to Turkish forms long enough for them to be expressed as /ɨ/ in the modern language. If they had not done so, we should not see /ɨ/ in the words in 0, since they do not contain any known conditioning environments.

Condition	Old Slavic	Romanian		
2	gǫsakŭ[1]	gânsac	/gɨnsak/	'gander'
	mǫdrŭ	mândri	/mɨndri/	'take pride'
3	krŭma[2]	cârmă	/kɨrmə/	'helm'
	stlŭpŭ	stâlp	/stɨlp/	'pillar'
	rysŭ	râs	/rɨs/	'lynx'
none	pyşanŭ	pâşen	/pɨʃen/	'haughty'
	bŭtŭ	bâtă	/bɨtə/	'stick, club'
	chudŭ	hâd	/hɨd/	'hideous'

Table 4: /ɨ/ from Slavic loans

Condition	Turkish	Romanian		
none	satır[3]	satâr	/satɨr/	'chopper, cleaver'
	gıdıklamak	gâdila	/gɨdila/	'tickle'
	kılıç	călâci	/kəlitʃ/	'sword'
	telhıs	talhâs	/talhɨs/	'functionary's report'

Table 5: /ɨ/ from Turkish loans

Besides the three sources presented briefly here, we find /ɨ/ in Romanian words that come from many other languages, including the modern Slavic languages Bulgarian, Polish and Serbian; Hungarian, and Greek. Among the words Romanian has borrowed it can be unclear whether /ɨ/ is a result of the conditioning environments and vowel quality within the donor languages, or whether they are a result of phonological changes internal to Romanian.

I argue that Romanian /ɨ/ began in allophony with /ə/, as is seen in native vocabulary, and gradually expanded to phonemic status with help from borrowings, especially from Old Slavic and Turkish. We see evidence for this in the increasing number of words in which /ɨ/ is phonologically unconditioned in loanwords. Historically, /ə/ and /ɨ/ were first attested as separate phonemes in the 16th Century, but the two remain separated by only a few minimal pairs: their distribution remains nearly complementary.

Having outlined the etymological history of /ɨ/, we now turn to its synchronic role within the current phonological system.

---

[1] In Slavic, < ǫ > represents a nasalized vowel, which Romanian treated as a /VN/ sequence.

[2] Attested Old Church Slavonic. In *Common Slavic, the positions of the liquid and vowel (jer) were reversed.

[3] In Turkish, [ɨ] is transcribed as < ı >.

3 The role of /ɨ/ in modern Romanian

A picture of the modern Romanian language, seen here, lets us view the role of /ɨ/ within the current phonological system. We examine the footprint of the aforementioned borrowings, and find that beyond the results of the phonological processes that caused raising and backing of Latin vowels to /ɨ/, this vowel has not spread to large portions of the Romanian lexicon. Very few Romanian words contain /ɨ/, and its appearance is often predictable based on its formerly allophonic relationship with /ə/: few minimal pairs distinguish the two phonemes, yet they are contrastive.

/ə/			/ɨ/		
rău	/rəw/	'bad'	râu	/rɨw/	'river'
ţări	/tsərʲ/	'lands'	ţâri	/tsɨrʲ/	'a little'
văr	/vər/	'cousin'	vâr	/vɨr/	'I thrust'

Table 6: Minimal pairs between /ə/ and /ɨ/

A type frequency analysis, in §4, confirms these observations: while /ɨ/ *can* appear in nearly any phonological environment, it tends to appear in pre-nasal position and near /r/, consistent with its allophonic origins. The distribution of /ɨ/ reflects the phonological footprint of its origins.

4 Phoneme Type Frequency in Romanian

This section presents a type-frequency analysis of wordforms, which are specific phonological forms (Levelt 1983). The types within a text are abstract categories defined for the purpose of the analysis: here, each of the phonemes of Romanian represents a type. Within a particular type, a token is a specific instance of that type; for example, Romanian *casa* /kasa/ contains three phoneme types, /k/, /s/ and /a/, and has two tokens of phoneme type /a/. Word forms are also known as lexemes, which contrast with lemmas – semantically and syntactically defined lexical entries. The relationship between word forms and lemmas is not one-to-one: a particular phonological form may correspond to multiple lexemes, and a lexeme may have multiple word forms, as in English *a – an*, and *the*, which is pronounced in at least two ways, [ðə] and [ði] (Jurafsky et al. 2002: 3).

The word forms for Romanian appear in a word list (Vladutu 2009) created for journalistic spell-checking purposes. The list's characteristics appear in 0. For the remainder of this discussion, frequencies of particular segment types are given as percentages; all percentages of vocalic segments are calculated relative to the number of total vowel characters, and all percentages of consonantal segments are calculated relative to the number of total consonant characters.

Total words	88,580
Total vowel characters	368,970
Total consonant characters	419,149
Total characters analyzed	788,119
Total characters	*788,157*[4]

Table 7: Word list statistics

4.1 Relative Frequencies of Vowels

First, I show type frequencies of vowels. In Figure 1, the Y-axis shows a count of the total instances of a particular vowel, represented as a percentage of the total number of vowel characters in the word list.

---

[4] This calculation includes a handful of characters that are not part of the Romanian alphabet, including numerals, vowels with accent marks, etc.
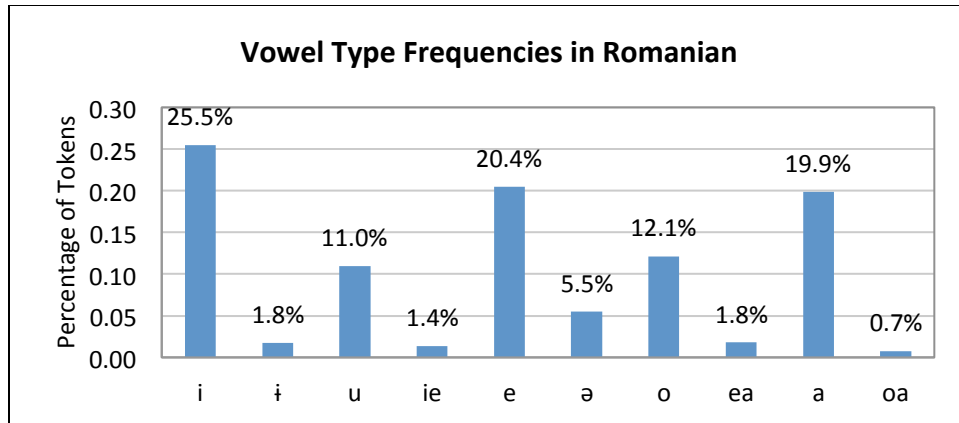
Figure 1: Romanian vowel frequencies

In 0, we see that <i> is the most frequent vowel in Romanian, accounting for 25% of vowel characters. While <i> represents the phoneme /i/, it does not always surface as [i]; it also appears as a glide, or as a word-final morphological marker, in the masculine plural and some verb conjugations. The vowels <a> and <e> are also very frequent in Romanian; together, these three vowels account for 70.1% of the vowels in our word list. <a> and <e> and are frequent morphological markers: among other things, <a> marks the definite form of many feminine singular nouns and a class of verbal infinitives, while <e> occurs within feminine plural endings and in verb forms. <o> accounts for 12.4% of the vowels analyzed here, followed by <u>, at 10.6%. We note that <o> might have higher token frequency: it is the feminine singular indefinite article, which would find frequent use in texts but is not seen here. On the other hand, <u> is part of the masculine singular definite marker /ul/, which is represented in the current word list.

Among the single phonemic vowels of Romanian, /ə/ and /ɨ/ are the least frequent: of all the vowels in this word list, /ə/ accounts for 5.3%, while /ɨ/ constitutes only 1.7% of the total vowel count. Both vowels occur in morphological inflections, although /ə/ has a more basic role than /ɨ/ in these patterns. /ə/ is the non-definite nominative/accusative ending for many feminine singular nouns, and often appears in our word list as an inflection; it is also frequent in verb paradigms.

On the other hand, /ɨ/ does not have a significant role in nominal or adjectival morphology – which all the other vowels do. It marks a subset of 4th conjugation verbs, and is the vowel in the gerund form of some verbs, such as *încercând* /intʃerkɨnd/ 'testing, trying'. From these facts, we can see that the low type frequency of /ɨ/ is not surprising, since its role in morphology is less than that of the other vowels.

The three diphthongs /jɛ/ <ie>, /ja/ <ea> and /wa/ <oa> also have very low type frequency. The latter two appear mainly in singular forms due to the effects of vowel metaphony, and all three appear only under stress, which greatly reduces the number of syllable nuclei they can fill. For reasons of space, I do not further address diphthongs here.

The type frequencies and morphological roles of Romanian's phonemic vowels indicate an uneven distribution of labor among the vowels. I show that in particular, the low type frequency of /ɨ/ correlates with its historical development. As we have seen, it did not spread much beyond the environments it occupied when it was an allophone. We find that although /ɨ/ can co-occur with nearly all the segments in the Romanian system, it still has characteristics of a phonologically-conditioned allophone. To complete this argument, however, we must consider not only type frequency among vowels, but also among consonants.

4.2 Relative Frequencies of Consonants

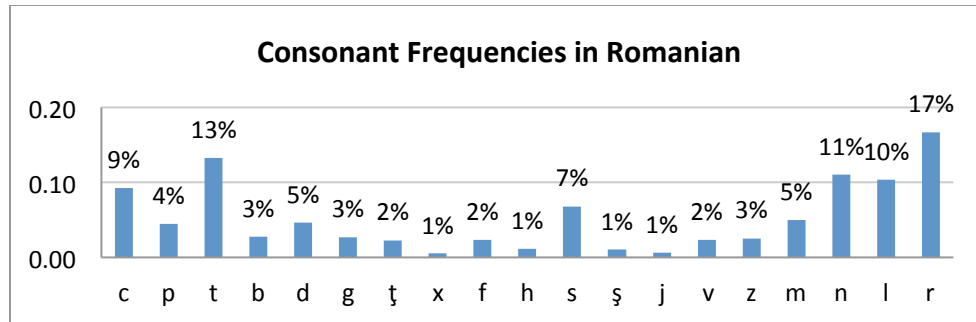0 shows the relative frequencies of consonants in Romanian.

Figure 2: Romanian consonant frequencies

In 0, we see that <r> is the most common of the Romanian consonants, making up 16.7% of the total. Close behind are <t>, <n> and <l>, which each make up more than 10% of the consonants in the word list. The segments <c> and <s> make up 9.3% and 6.8%, respectively, of the consonants; together, these five segments account for 67.4% of the consonants analyzed here, and no other consonant type makes up more than 5% of the total.

Knowing now which Romanian consonants are the most frequent types, we are better equipped to judge the frequencies with which each consonant type appears adjacent to a given vowel type. We should not be surprised to see /i/ appearing frequently near /r/ or /n/, for example, because all three of these segments are quite frequent in Romanian. On the other hand, if we find that a low-frequency type occurs frequently with another type, we may suspect that their co-occurrence is more than coincidental. This is the case for /ɨ/.

5 Romanian Vowels and their Following Segments

Here we compare the role of /ɨ/ to that of the other vowels by examining the segments that surround them. In the following figures, I show the array of segments that appears *after* each of Romanian's phonemic vowels, and I demonstrate that there is a strong relationship between /ɨ/ and a following nasal. No relationship of this magnitude exists between another vowel and a following segment, correlating with the historical phonological relationship between nasals and /ɨ/.

Figures 3–9 are ordered from most frequent to least frequent vowel type: /i/ is shown first, and /ɨ/ last.[5] In these figures, the X-axis shows the array of segments found after each vowel type, and the number atop each bar is a percentage, for example showing the percentage of /i/'s within the word list that are followed by a particular segment. The Y-axis shows a count of instances of each following segment, represented as a percentage of the total instances of /i/. In these figures, # stands for a word boundary.

In Figures 3–8, we see the distributions of segments following six Romanian vowels. For all of these except /o/ and /u/, the most common following environment is a word boundary, indicating the frequency with which these vowels are word-final. This is especially true for /ə/, which falls at word boundaries 45.9% of the time. These are mostly feminine nouns and adjectives, evidence for the morphological link between vowel type and frequency. Other morphological markers at word boundaries are /i/, in masculine plurals and infinitives; /e/, in feminine plurals; and /a/, in feminine definite forms and infinitives.

In 0 we see that nearly half the instances of /ə/ are word-final, but otherwise, no percentage in these figures reaches even 30%; and the highest percentages are all linked to the most common consonants of Romanian or to morphological facts. In other words, we have no evidence that phonological conditioning by a certain consonant affects the distribution of these vowels. In 0, we see that these tendencies do not hold for the set of segments that follow /ɨ/.

In 0, we see a picture strikingly different from the other six vowel phonemes. The vast majori-

---

[5] The diphthongs <ie>, <ea> and <oa>, are not analyzed here, and have been removed from the remaining figures and calculations. The motivation for this exclusion is that the diphthongs do not appear to compete with /ɨ/ for vowel slots.

ty of tokens of /ɨ/ — a total of 74% — precede /n/ in Romanian. Another 10% precede /m/; and 7.7% are followed by /r/. No other segment follows more than 2% of instances of /ɨ/. Of the 622 instances of /ɨm/, 93% belong to the sequence /ɨmC/; and of the 472 instances of /ɨr/, 86% belong to the sequence /ɨrC/, recalling the rules in 2.

   These three frequent following environments correspond to the phonological processes that gave rise to /ɨ/ in Romanian. This indicates that while /ɨ/ can appear in the vicinity of various consonants in Romanian, with the effect that a reliable rule about its distribution is not possible, there are very strong tendencies that help describe or predict the appearance of /ɨ/. In other words, the allophonic history of /ɨ/ has left a strong mark on the segment's distribution, and the use of /ɨ/ has not greatly expanded beyond its original conditioning environments.
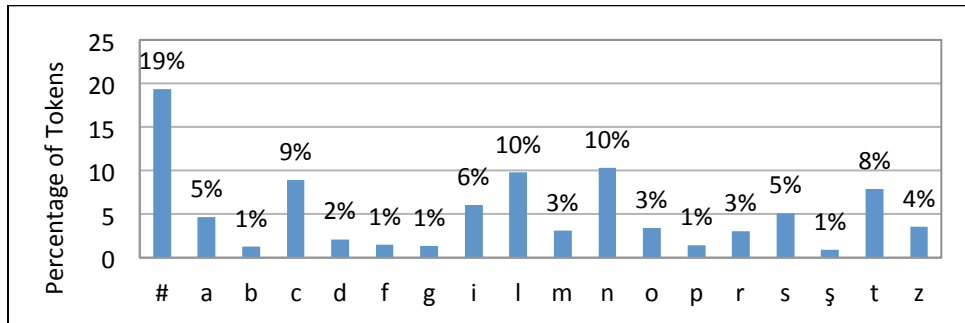
Figure 3: Segments following <i> in Romanian[6]
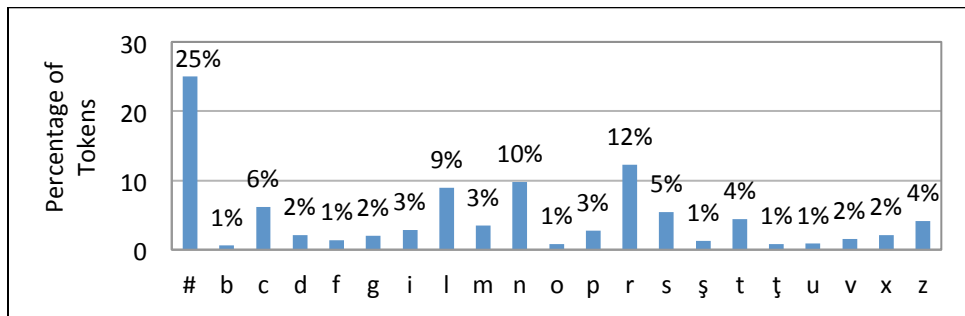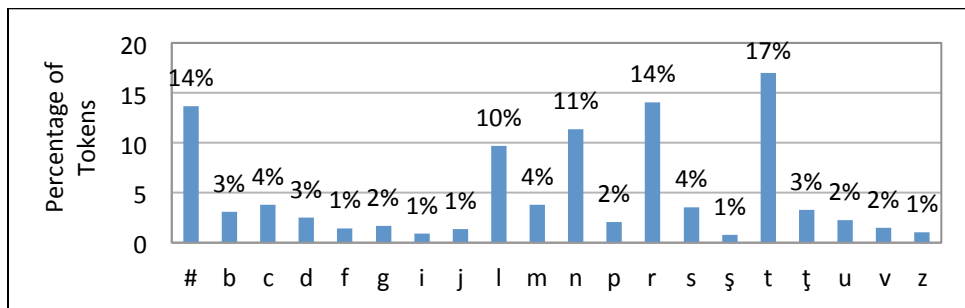

Figure 4: Segments following <e> in Romanian

Figure 5: Segments following <a> in Romanian

---

6 **NOTE**: In Figures 5–8, all segments that followed the vowel in less than 1% of cases were excluded from the graph, for illustration purposes.
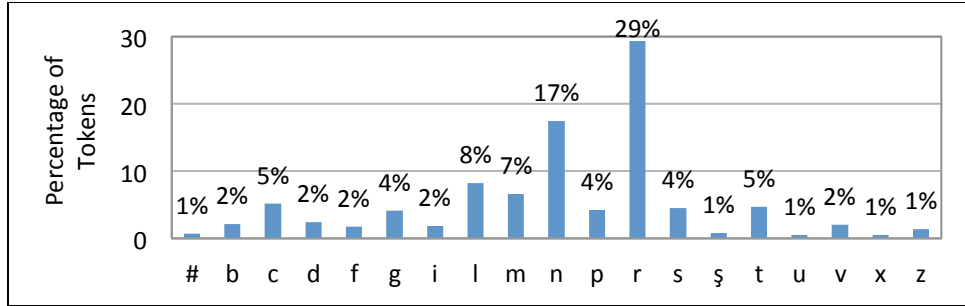
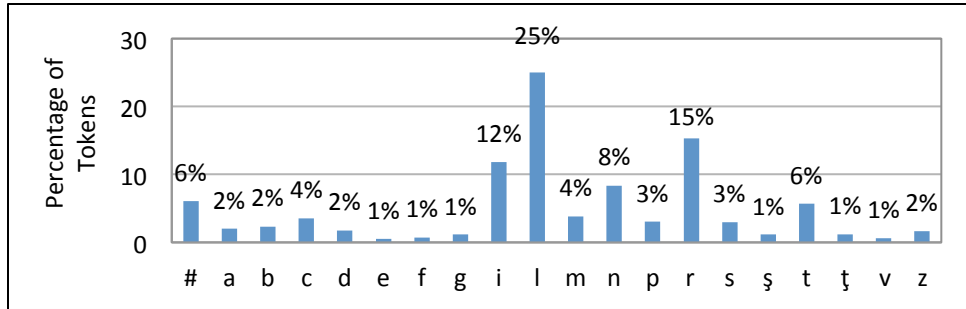Figure 6: Segments following <o> in Romanian
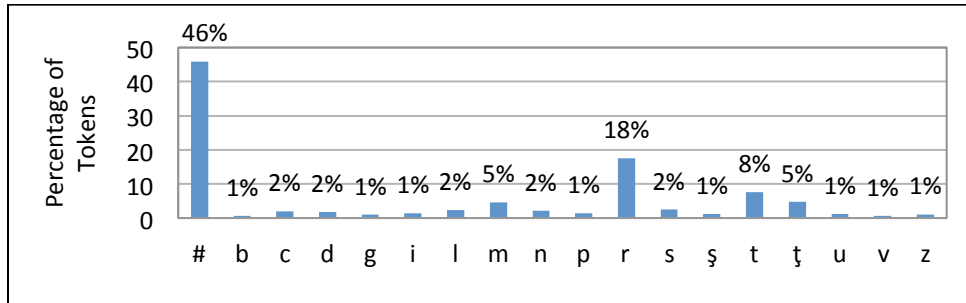

Figure 7: Segments following <u> in Romanian

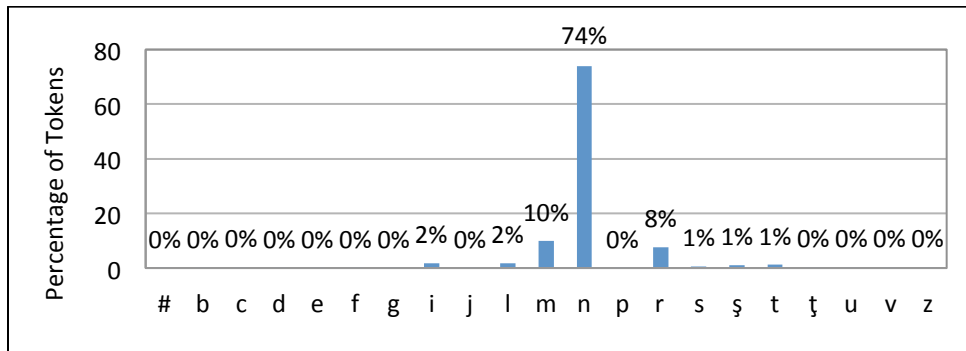Figure 8: Segments following /ə/ (transcribed <ă>) in Romanian

Figure 9: Segments following /ɨ/ (transcribed < â > or < î >) in Romanian

6 Frequency of Romanian Vowels before /n/

To fully picture the relationship between /ɨ/ and its surrounding environment, we can also look at the co-occurrence of /ɨ/ and /n/ from the point of view of the latter, which is a highly frequent consonant in Romanian. Since /ɨ/ is very infrequent in the language, we might not expect it to make up a large percentage of the preceding segments for /n/; however, if the two do tend to co-occur, this is further evidence that they lie in a phonologically-conditioned relationship.

I argue that this is the case, based on 0, which shows a comparison between type frequency of Romanian vowels across all environments (expected frequency), and pre-/n/ type frequency. If /ɨ/ is more frequent before /n/ than elsewhere, the evidence of an interaction between /ɨ/ and /n/ is strengthened. 0 shows that for several Romanian vowels, a comparison of overall type frequency with type frequency before /n/ finds that these vowels are less frequent before /n/ than they are overall. For /ɨ/ the opposite is true. In pre-/n/ position, /ɨ/ makes up 11% of the vowels instead of the 1.7% it occupies across all environments. Although /i/ is over 20 times more frequent than /ɨ/ in the language overall, it is only twice as likely to occur before /n/.
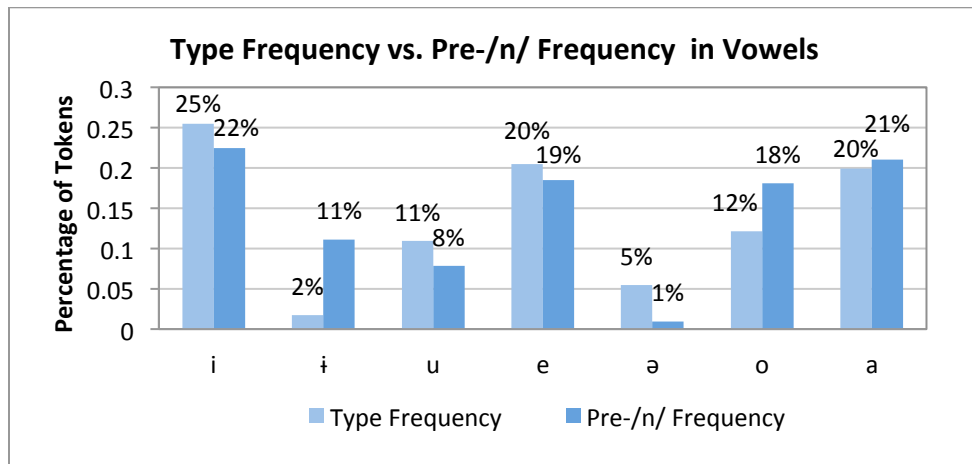

Figure 10: Comparison of vowels' overall type frequency and pre-nasal frequency

Finally, in Figure 11, we see the ratio of each vowel's pre-/n/ frequency to type frequency. While most vowels' ratios are near 1, the type frequency of /ɨ/ increases more than six fold when /n/ is the following segment. In sharp contrast, the ratio for /ə/ is very small, at 0.18, showing the former allophones' complementary distributions. This reinforces the evidence that the distribution of /ɨ/ results from phonological conditioning.
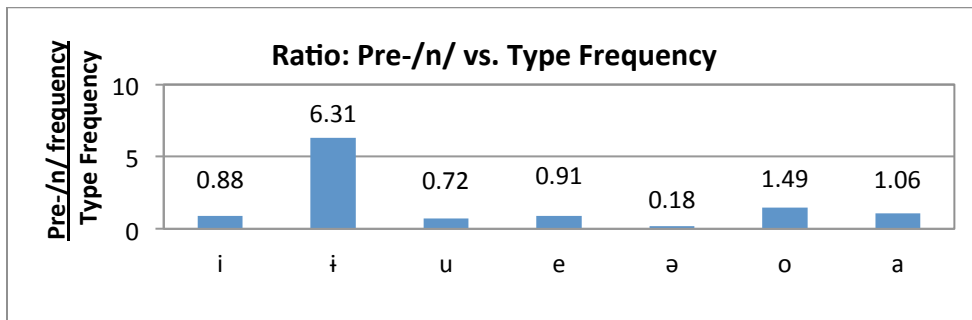

Figure 11: Ratios of pre-/n/ to overall type frequency in Romanian vowels

7 Summary

Historically, /ɨ/ was conditioned in Romanian by several factors, the greatest of which were a) a following nasal, or b) a following /rC/. Among native words, /ɨ/ was allophonic, in a complementary distribution with /ə/ in particular. In modern Romanian, /ɨ/ is phonemic; loanwords helped

cement its status as underlying.

We find that /ɨ/ has very low type frequency, making up less than 2% of all vowels in Romanian word forms. However, /ɨ/ appears in high token frequency words, like *în* /ɨn/ 'in' and *îmi* /ɨmʲ/ 'to me'; this could help account for its continued phonemic status.

While /ɨ/ is a phoneme, its distribution is highly constrained as a result of its history. The phonologically-conditioned footprint of these historical processes is in fact seen in the frequencies with which other consonants appear around /ɨ/, specifically after the vowel: nasals and /r/ make up more than 90% of the following environments for this vowel. Co-occurrences of this magnitude are not seen elsewhere, in the data from the other six vowels of Romanian.

Using a type frequency analysis, we not only confirm the phonological conditioning that originally brought /ɨ/ into Romanian; we also see that the allophonic history of /ɨ/ has left a strong mark on the segment's distribution, and that the use of /ɨ/ has not greatly expanded beyond its original conditioning environments.

References

Coteanu, Ion. 1981. *Structura şi evoluţia limbii române (de la origini pînă la 1860)*. Bucharest: Editura Academiei Republicii Socialiste România.

Hall, Robert. 1974. *External History of the Romance Languages*. New York: American Elsevier Publishing Company.

Jurafsky, Daniel, Alan Bell and Cynthia Girand. 2002. The role of the lemma in form variation. In *Laboratory Phonology 7*, ed. C. Gussenhoven and N. Warner, 3–34. New York: Mouton de Gruyter.

Levelt, W. J. M. 1983. *Speaking: From Intention to Articulation*. Cambridge, MA: MIT Press.

Mallinson, G. 1988. Romanian. In *The Romance Languages*, ed. M. Harris and N. Vincent. London & Sidney: Croom Helm.

Petrovici, Émile. 1958. Interpénétration d'une phonologie slave et d'une morphologie roumane. *Mélanges linguistiques*, ed. I. Iordan, É. Petrovici, and A. Rosetti. Bucharest: Academia Republicii Populare Romîne.

Petrucci, Peter R. 1999. *Slavic Features in the History of Rumanian*. Munich: Lincom Europa.

Renwick, Margaret E. L. 2009. *On the origins of /ɨ/ in Romanian*. Unpublished ms., Cornell University.

Rosetti, Alexandru. 1986. *Istoria limbii române I: de la origini pînă la începtul secolului al XVII-lea*. Bucharest: Editura Ştinţifică şi Enciclopedică.

Sala, Marius. 1976. *Contributions à la phonétique historique du roumain*. Paris: Klincksieck.

Schulte, Kim. 2005. Vowel Centralization in Romanian Verbs of Slavic Origin: Deliberate Exploitation of an Indigenous Sound Change? In *Romance Languages and Linguistic Theory 2003* [Current Issues in Linguistic Theory 270], ed. H. Jacobs and T. Geerts, 378–394. Amsterdam: John Benjamins Publishing Company.

Vasiliu, Emanuel. 1968. *Fonologia istorică a dialectelor dacoromâne*. Bucharest: Editura Academiei Republicii Socialiste România.

Vladutu, Daniel. 2009. Romanian word list for journalistic applications. Personal communication, February 6, 2009.

Department of Linguistics
Cornell University
Ithaca NY 14850
mer56@cornell.edu