McGill University

Final Paper

Notable Phonological Processes and Opacity in Bulgarian

Presented by

Kaloyan Kanev

Presented to

Dr. Andrei Munteanu

In the context of

LING 531

Phonology 2

Winter 2025

April 16th, 2025

This paper serves as an improved, synchronic phonological description of modern Bulgarian, describing and illustrating phonological processes that are commonly attested, but also highlighting processes that aren't mentioned in the literature. It explores how processes interact, sometimes resulting in opacity. In a first part, I will start by examining the various phonological processes observed in Bulgarian, including common but systematic alternations observable in casual speech. In a second part, I will focus on the interaction between Voicing Assimilation and Word-Final Devoicing. Depending on their ordering, these processes can potentially interact in either a feeding or counterfeeding relationship, which could result in opaque surface forms. I will discuss how such opacity is theoretically possible and demonstrate how it can be observed in a vacuum, while highlighting that in practice, no concrete instances of this interaction—whether resulting in opacity or not—are attested in Bulgarian. In a third part, I will turn my attention to the interaction between Nasalization and Nasal Deletion, which I will illustrate using relevant examples. Here, the ordering of the processes does result in attested opaque surface forms. I will also analyze this interaction through the lens of autosegmental phonology, showing how opacity is mostly a product of our analysis. Throughout the paper, I will also use Bulgarian orthography as a guiding tool in my analysis. Given the language's relatively transparent orthographic system, I argue that orthography plays a crucial role in identifying and understanding phonological patterns, and I argue that it allows speakers a greater awareness of the underlying forms of utterances. This approach differs from usual phonological analyses, as discussions on orthography are often avoided. The aim of this paper is to improve our understanding of the variations observed in spoken Bulgarian, and hypothesize on the possible implications this has for the language in the future.

1. Bulgarian Language

Bulgarian is a language spoken by approximately eight million people today. As a member of the Slavic language family, it shares many linguistic features with related languages such as Russian. Over time, it has also incorporated a significant amount of borrowed vocabulary from neighbouring languages. Today, Bulgarian serves as the official language of Bulgaria, and its standardized form is known as the Bulgarian Literary Language—a formal variety defined and regulated through its orthography.

In an increasingly literate world, proficiency in the standard variety of a language is often associated with high literacy. While this connection is not necessarily causal, it remains important to consider orthographic representations in our analyses, especially in languages with relatively transparent orthographies like Bulgarian. Speakers often perceive orthography as a faithful reflection of the underlying linguistic representation. In the case of Bulgarian, this transparency supports the prescription and enforcement of a standard variety through its writing system.

Most Bulgarian letters correspond directly to individual phonemes, with only a few exceptions. This phoneme-to-grapheme alignment is particularly useful for linguists, as it provides insights into the underlying representations of words. Since the orthography reflects the prescribed standard, it also influences how speakers articulate words in careful speech. They often adhere closely to the written form, sometimes even emphasizing phonemes that would otherwise undergo phonological changes in casual or rapid speech. This phonetic transparency can, however, be a double-edged sword. It is often observed to lead speakers to make orthographic errors when attempting to write based on how words sound in everyday speech. In this paper, we will highlight such instances, particularly where they intersect with relevant phonological processes.

Bulgarian is a language I speak, write, and used to think in. This native familiarity comes with clear advantages—most notably, an intuitive awareness of the sounds that surface during speech production. It also enables me to readily identify relevant examples of phonological processes, many of which will be explored in this paper. At the same time, native knowledge can be a disadvantage. Some linguistic phenomena may not stand out to a native speaker due to the very familiarity that makes the language feel natural. When we think in a language, we're not always thinking *about* the language.

2. Bulgarian Phonology

Bulgarian exhibits a range of interesting phonological processes, many of which are shared with other Slavic languages, while others are unique to Bulgarian itself. In this section, we will formalize these processes and illustrate them with relevant examples. The following section will then focus on how certain pairs of these processes interact.

Sociolinguists often distinguish between careful and casual speech. Due to the relatively transparent nature of Bulgarian orthography, speakers will often rely on the written form as a guide for pronunciation. One notable consequence of this is the tendency for speakers to emphasize underlying forms in careful speech—even when these forms differ from the phonological surface realization. For instance, a common phonological process in Bulgarian is word-final devoicing. As a result, the word for "scallion" is underlyingly /praz/ but surfaces as [pras]. However, when a speaker wishes to emphasize the word (for example, to correct someone or draw attention to it) they will produce it as [praz], aligning with the orthographic form and underlying representation, which contains the letter **з** [z].

This phenomenon supports the idea that alternations observed in casual Bulgarian speech should be analyzed as phonological rather than purely sociolinguistic or dialectal. Bulgarian speakers appear to have a heightened awareness of the underlying forms of words, largely due to the transparency of the orthography. Additionally, the standardization of the language through orthography makes it an explicit

guideline for speech. Therefore, I argue that recurring surface-level allophonic variations are phonologically driven, especially when they consistently occur in the same phonological environments. Alternations that appear more sporadically or irregularly, on the other hand, will be treated as sociolinguistically motivated. The analyzed processes have been observed to occur in speakers of Standard Bulgarian, including myself, my family and friends of family. Any dialectal variations will not be featured as part of the relevant phonological processes in this paper, and if I am to use examples of dialectal variations, they will be explicitly labeled as such.

2.1 Unstressed Vowel Raising

Bulgarian has 6 phonemic vowels: /i, ɛ, ɤ, a, u, ɔ/. All of these vowels appear underlyingly:

мир /mir/ “peace” **път** /pɤt/ “road” **тук** /tuk/ “here”

ред /rɛd/ “order” **над** /nad/ “above” **рог** /rɔg/ “horn” (of an animal)

Some can also appear as allophones of other vowels: /ɛ, a, ɔ/ can become [i, ɤ, u] in certain environments. A very commonly observed process is the raising of unstressed vowels, most notably from [ɔ] to [u], and can be often heard word initially or finally. A vowel that is underlyingly mid or low will be raised to high or mid respectively when in an unstressed position. (Scatton, 1993)

много	/ˈmnɔgɔ/ → [ˈmnɔgu]	“many”
зелено	/zɛˈɬɛnɔ/ → [ziˈɬɛnu]	“green”
отивам	/oˈtivam/ → [uˈtivɤm]	“I am going”

This process can create homophony and potential confusion during speech.

очи	/ɔˈtʃi/ → [uˈtʃi]	“eyes”
учи	/ˈutʃi/	“learn”

As we have previously mentioned, Bulgarian orthography is relatively phonetically transparent, which allows speakers and writers of the language to infer the orthography of a word using its sound. This vowel raising process is therefore reflected in a common spelling mistake young people make, putting ***у*** [u] word finally or word-initially when it should instead be ***о*** [ɔ]. The confusion between ***а*** [a] and ***ъ*** [ɤ] is also a commonly seen phenomenon, although the same is not observed for the raising of ***е*** [ɛ] to **и** [i]. The latter tend to be spelled correctly. This further highlights the important connection between orthography and the phonology of Bulgarian.

2.2 Word Final Devoicing

Bulgarian, like most other Slavic languages, features devoicing of voiced consonants in word final position. (Scatton, 1993) These consonants easily resurface when a speaker wishes to emphasize them, showing an inherent knowledge of the underlying form of the word.

какъв	/kakɤv/ → [kakɤf]	"of what kind"
гърмеж	/gɤrmeʒ/ → [gɤrmeʃ]	"thunder"
праз	/praz/ → [pras]	"scallion"

When suffixed, the underlying voiced consonant surfaces.

праз	/praz/ → [pras]	"scallion"
празове	/prazove/ → [prazuve]	"scallions"
град	/grad/ → [grat]	"city"
градове	/gradɔvɛ/ → [graduvɛ]	"cities"

Devoicing is blocked in prepositions.

град ли?	[grat ɬi]	"city?"
пред лицето	[prɛd ɬitsɛtɔ]	"in front of the face"

2.3 Voicing Assimilation

Assimilation occurs in consonant clusters. Voiced consonants devoice if followed by voiceless consonants, and voiceless consonants voice if followed by voiced consonants:

изток	/iztɔk/ → [istɔk]	"east"
изход	/izhɔd/ → [ishɔd]	"exit"
сграда	/sgrada/ → [zgrada]	"building"

An interesting word to note here is the word **изграждане** /izgraʒdanɛ/ "construction". With the last word for "building" [zgrada], we could be led to believe that /izgraʒdane/ is a morphologically related form with a reanalyzed /z/. However, the prefix here is simply **из-** [iz-].

Voicing assimilation also occurs across word boundaries (sometimes called sandhi).

над полето	/nad pɔɬɛtɔ/ → [nat pɔɬɛtɔ]	"above the field"
от гората	/ɔt gɔrata/ → [ɔd gɔrata]	"from the forest"

Voiced fricatives appear to behave differently with regard to Voicing Assimilation. When /v/ is preceded by a consonant, it is /v/ that devoices instead of the preceding consonant:

какво	/kakvɔ/ → [kakfɔ]	"what"

However, Voicing Assimilation with /v/ still occurs across word boundaries:

от вдовица	/ot vdɔvitsata/ → [od vduvitsɤtɤ]	"from the widow"	(Scatton, 1993)

2.4 Nasalization

In the literature, this is often not treated as a process of its own, probably because it is so widespread cross-linguistically and arguably a purely articulatory constraint. That being said, we need it for our analysis of an opaque interaction between it and n-elision.

Nasalization occurs with vowels followed by a nasal, specifically /n/, that is itself followed by a consonant. Otherwise, there is no nasalization, or at least none that is salient enough to be noted (because we are bound to have *some* nasalization in pre-nasal vowels due to coarticulation).

кон	/kɔn/ → [kɔn]	"horse"
конски	/kɔnski/ → *[kɔ̃nski]	"of horse type"
тон	/tɔn/ → [tɔn]	"tone"
тонколона	/tɔnkɔlɔna/ → *[tɔ̃nkulɔna]	"speaker"

The forms with nasalization are marked as unattested is because of the presence of the nasal. There is another process in Bulgarian that causes the /n/ to delete between vowels and consonants.

However, the same phenomenon cannot be observed with the nasal /m/:

том	/tɔm/ → [tɔm]	"tome"
компот	/kɔmpɔt/ → [kumpɔt]	"compote"

The example for the word "compote" shows no nasalization of the preceding vowel, but also doesn't exhibit any deletion of the nasal consonant. That being said, most words with [-ɔmC-] sequences —including "compote"—in Bulgarian are not of Proto-Slavic origin and are often borrowings from Latin ("complex" /kɔmplɛks/, "company" /kɔmpanja/, "computer" /kɔmpjutɤr/). Nonetheless, borrowed words still should undergo the receiving language's sound changes, so this gap indicates that vowel nasalization in Bulgarian only occurs before /n/, and not all nasal, which would include /m/.

2.*5 N elision*

As mentioned, the underlying /n/ is deleted after nasalizing. The vowel then becomes long.

конски	/kɔnski/ → [kɔ̃ːski]	"characteristic of/belonging to a horse"
тонколона	/tɔnkɔlɔna/ → [tɔ̃ːkulɔna]	"speaker"

2.6 Dark L

A relatively new phenomenon in Bulgarian is the lenition of the phonemic Dark L [ɬ] to [w]. (Padareva & Mitsova, 2014) Dark L is characterized by its dental place of articulation, where the tip of the tongue touches the back of the teeth. Bulgarian ***л*** [ɬ] is always dark, however it alternates between a dental and alveolar place of articulation in certain situations:

колело	/kɔɬɛɬɔ/ → [kɔlɛɬɔ]	"bicycle"

Dark L itself also varies, although slightly, in its dental contact. Some instances of [ɬ] will have an obvious contact between the tip of the tongue and the back of the teeth (marked here with underline):

белтък /bɛɬtɤk/ → [bɛ̲ɬtɤk] "egg white"

In other situations, there is no contact, but it remains Dark L, as a dental contact is not necessary.

булка /bułka/ → [bułkɤ]

More importantly, newer generations have been leniting Dark-L to [w], which is not a phoneme of the standard variety of Bulgarian. It is observed in some regional dialects, however it has been spreading throughout urban areas as well, most prominently among young people. Research on this is still scarce, but it also shows that people who produce [w] not only cannot produce [ɬ], but they also can't perceive the difference between [w] and [ɬ] in others' speech. (Padareva & Mitsova, 2014) A proposed hypothesis for this is borrowings from other languages with /w/, and since it's easier to produce than [ɬ], it is favoured and adopted instead of it. Other languages like Polish have also undergone this change and are now fully devoid of a [ɬ], instead producing [w]. Bulgarian might also be heading in that direction.

2.7 Coronal Stop Deletion

Coronal segments /t/ and /d/ delete, specifically between consonants and nasals. (Scatton, 1993) This process often occurs when speaking rapidly, and the stops can be brought back in careful speech.

вестник /vɛstnik/ → [vɛsnik] "newspaper"

издниквам /izdnikvam/ → [iznikvam] "I emerge"

Many speakers tend to misspell words where coronal stop deletion occurs, often omitting the phonologically deleted /t/ or /d/, because it is absent in their production. Speakers having surface knowledge of how a word is pronounced, but not it's etymology or morphological composition (e.g., вест-ник, из-дниквам), results in production-true transcriptions.

2.8 Intervocalic Voiced Fricative Deletion

Voiced fricatives /v/ delete intervocalically. This is most commonly observed in casual speech. It is sometimes argued that this is dialectal, however it is regularly observed in literate and educated speakers of the standard variety of Bulgarian. The motivation appears to be simplification of speech.

правя	/pravja/ → [praː]	"I'm doing"
правиш	/praviʃ/ → [praiʃ]	"you're doing"
трябва	/trjabva/ → [trjaa]	"needs"
хубаво	/ˈhubavɔ/ → [hubau] → [hubu]	"good"

In the final example, the intermediate form [hubau] has a diphthong in an open syllable. There seems to be a constraint against this kind of structure, turning the diphthong into a monophthong. Also note the unstressed vowel raising of /ɔ/ to [u], not crucially ordered in this example.

In monosyllabic words, the deletion is blocked and, instead, syncope is observed.

това /tova/ → [tva] → [tfa]

Also note the transparently interacting processes: syncope feeds voicing assimilation.

One hypothesis for such a variation is to preserve maximum information in words that have little information to begin with (i.e., monosyllabic words). It would allow listeners to infer the whole word through its consonants. Another hypothesis is related to the previously mentioned constraint against open syllable diphthongs. Deleting the /v/ would result in an open diphthong, so instead a vowel is syncopated to prevent it, leaving [v]. Below is a parallel OT account of these hypotheses:

Constraints:

*VvV	Assign a violation for every intervocalic /v/
ObsMaxMonoSyll	Do not delete obstruents in monosyllabic words
*OpenDiph	Assign a violation for every diphthong in an open syllable
Agree	Adjacent obstruents must agree in voicing
Max	Assign a violation for every segment in the input absent in the output

Maximal Consonants Hypothesis

/tɔva/	*VvV	ObsMaxMonoSyll	Agree	Max
tɔva	*!			
tɔa		*!		*
tfa				*
tva			*!	*

Open Diphthong Hypothesis

/tɔva/	*VvV	*OpenDiph	Agree	Max
tɔva	*!			
tɔa		*!		*
tfa				*
tva			*!	*

Both of these hypotheses essentially accomplish the same thing, so deciding which one is more appealing can fall on which constraint is more realistic, or least ad-hoc. ObsMax is attested in OT accounts of other languages, while there is little evidence for a constraint against open diphthongs. However, the latter constraint would be in line with the evidence in /hubavɔ/ → [hubau] → [hubu].

Another interesting example of /v/ deletion is the following:

остави /ɔstaˈvi/ → [ustɤˈvi] → [‘ɔstɤj] “leave (it)”

Both underlyingly and on the surface, before any /v/ deletion occurs, we observe stress on the final syllable. However, when /v/ is deleted intervocalically, stress falls onto the first syllable. This is difficult to account for phonologically. Since many Bulgarian dialects have variations in stress assignment, we will posit that this instance of stress moving after /v/ deletion is dialectal. Note that the word initial /ɔ/ remains (or goes back to) mid [ɔ], because it is now stressed.

This process is relevant to point out as it seems to be missing from the literature, even thought it is widespread among speakers of Bulgarian, including of the standard variety. Further analysis on the relationship between v-deletion and stress—or just, overall simplification and stress—is relevant, since these are patterns standard speakers of Bulgarian exhibit.

3. Interactions

3.1 Word Final Devoicing and Voicing Assimilation

We have previously described and illustrated these two processes observed in Bulgarian, but also other Slavic languages. Recall some of the examples of Word Final Devoicing and Voicing Assimilation:

какъв	/kakɤ**v**/ → [kakɤ**f**]	“of what kind”
гърмеж	/gɤrme**ʒ**/ → [gɤrme**ʃ**]	“thunder”
праз	/pra**z**/ → [pra**s**]	“scallion”
изток	/i**z**tɔk/ → [i**s**tɔk]	“east”
изход	/i**zh**ɔd/ → [i**sh**ɔd]	“exit”
сграда	/**sg**rada/ → [**zg**rada]	“building”

These processes are formalized into the following rules to better observe their interaction:

Word Final Devoicing C → [-voice] / __#

Voicing Assimilation C → [αvoice] / __[αvoice]

Ordered in the above way, Devoicing feeds Assimilation. By devoicing a word-final segment, it could potentially cause a preceding, voiced segment to devoice as well. Ordered in the opposite way, this would be a counterbleeding relationship. What we would need for this interaction to occur is a word-final, voiced consonant cluster. This process can be observed in Russian, where such clusters are present:

UR:	**мозг**	/mozg/
Devoicing:		k
Assimilation:		s
SR:		[mosk]

This process, however, is not observed in Bulgarian, because Bulgarian doesn't have such clusters, at least not anymore. In fact, "brain" in Proto-Slavic (from which both Bulgarian and Russian descend) is ***мозгъ**. Bulgarian repaired this word through metathesis: **мозък**; while Russian removed the final vowel and kept the voiced consonant cluster intact: **мозг**. This is an instance of a systematic gap, where these processes could've interacted in theory, but there are no environments that allow for it.

It is relevant to mention that few words with final voiced consonant clusters can, in fact, be observed in Bulgarian. One such example is **вожд** /voʒd/ "leader". This word does also surface as [voʃt], as predicted by the rules and their ordering. Why this word wasn't repaired in a similar way as for "brain" is unknown. In Russian, the same word is **вождь** /voʒdʲ/, which surfaces as [voʃtʲ]. In Bulgarian, the lack of a phonemic palatalization like in Russian probably allows for the final stop to surface as unreleased, and the [ʒ] often still appears on the surfaces voiced. For now, there isn't a valid explanation for this gap.

The main takeaway is that Word Final Deletion and Voicing Assimilation have the potential to interact, and they do so in other Slavic languages like Russian. This allows us to infer that they are ordered in a way that results in feeding. However, Bulgarian exhibits a systematic gap, as it doesn't have environments that allow for this interaction to occur, probably due to other, historical repair processes that have already dealt with the unwanted word-final voiced consonant clusters.

3.2 Nasalization and N-Elision

Recall the examples of Nasalization and N-Elision:

кон	/kɔn/ → [kɔn]	"horse"
конски	/kɔnski/ → [kɔ̃ːski]	"of horse type"
тон	/tɔn/ → [tɔn]	"tone"
тонколона	/tɔnkɔlɔna/ → [tɔ̃ːkulɔna]	"speaker" (tone-column)

These processes can be formalized using the following two rules:

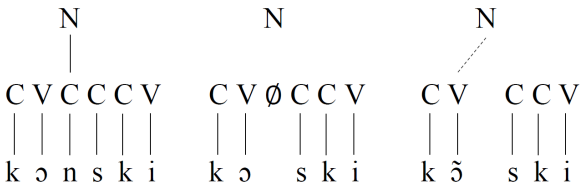
Nasalization: V → [+nasal] / __nC

N-Elision: n → ∅ / __C

Ordered in the above way, Nasalization appears to have overapplied. N-Elision counterbleeds Nasalization, as it would've removed the environment for Nasalization. This is counterbleeding opacity.

UR:	kɔnski
Nasalization:	ɔ̃
N-Elision:	∅
SR:	kɔ̃ːski

However, opacity is a product of our analysis, it is a result of the choice of framework that we use to explain data. Opacity is sometimes an indication that something has gone awry, and it is often a motivation for linguists to change the framework they use to obtain a more appealing explanation of ongoing processes. An autosegmental approach is appropriate to describe this phenomenon. There would be 3 tiers: the segmental tier, the CV tier, and the Nasal tier. This allows for the nasal feature to spread to the adjacent vowel when it's segmental or C tier is deleted. This is represented below:



The first representation is the underlying one, where a nasal feature is attached to the /n/ in /kɔnski/. On the second representation, the /n/ is deleted through the N-Elision process. The nasal feature remains, and it relinks to the preceding vowel tier, nasalizing the /ɔ/. With this, we are able to provide a more appealing explanation for the derivation of [kɔ̃ːski], without running into the issues of overapplication opacity. Such an analysis is particularly interesting, for it entails that Bulgarian Literary Language does not have any instances of phonological opacity.

Conclusion

There seems to be a growing trend in the spoken language to simplify speech as much as possible, while still maintaining enough information to allow for successful interpretation on the receiving end of an utterance. This isn't uncommon in languages and, in fact, forms that are a result of such a push towards simplification eventually fossilize into the standard variety of the language. Examples of this are many, including from English: it lost its cases, "art" became "are", "goeth" became "go", etc. Previously, we also mentioned the loss of Dark L in polish and its replacement with [w], and how Bulgarian might be undergoing this same change. Polish [w] used to be considered a poor use of language, used by the lower classes. The Bulgarian's public perception of this shift towards the articulatorily simpler [w] is also relatively negative, often associating it with regional non-standard dialects, or even teenage carelessness. The other variations that we've outlined in this paper are also proscribed forms of speech, and often regarded as dialectal by linguists. Today, these—often phonologically motivated—simplifications of the standard language might become the new standard language of tomorrow. Throughout this paper, we have also argued that orthography plays an important role in languages like Bulgarian, where graphemes and phonemes often align. Speakers of Bulgarian have an increasingly widespread understanding of the

written language, as this is also how it has been historically standardized. In fact, it's in the name: Bulgarian Literary Language. Therefore, it is not unreasonable to posit that speakers of the standard Bulgarian variety consciously rely on the orthographic representations of words to infer their underlying sounds. This alignment plays a role in the opposite direction as well, where speakers of Bulgarian, being aware of this alignment, will often make orthographic mistakes by trying to transcribe what they produce, not aware of the phonological processes that their productions undergo. In fact, throughout this paper, I have outlined some of the most important processes that can be observed in speakers of the standard variant. I have also highlighted the various interactions—both potential and attested—between some of these processes. I have pointed out an instance of opacity, while also proposing an analysis that avoids it, since opacity is just a product of our analysis, and potentially an indication that *we* are doing something wrong, not that there is something wrong.

Below are other instances of the push towards simplification that can be observed in Bulgarian. This list is simply for further curiosity, as it is difficult to account for all observed changes with a specific, regular, recurring phonological process. Most of these are a combination of multiple processes or might simply be dialectal—although I've previously highlighted the problem with such an easy explanation for variation. These variations are therefore relevant to point out as they are rarely—if ever—mentioned in the literature. They have been extracted from my own knowledge of the language, and the production of the speakers of standard Bulgarian in my entourage.

сега	/sɛga/ → [sɛa] → [sja]	"now"
хайде	/hajde/ → [aj]	"come on"
хайде сега	/hajdɛ sɛga/ → [aj sja]	"come on now"
какво	/kakvɔ/ → [kɤkfɔ] → [kfɔ] → [kɔ]	"what" neut.
какъв	/kakɤv/ → [kɤkɤf] → [kɤf]	"what" masc.
какво от това	/kakvɔ ɔt tɔva/ → [kɔ u(t) tfa]	"what about it"
всичко	/vsitʃko/ → [fsitʃko] → [sitʃko]	"everything"
трябва да се направи	/trjabva da se napravi/ → [trjaa d(a) s nɤprai]	"it has to be done"

References

Padareva, G. I., & Mitsova, S. (2014). Is Bulgarian Language Losing Its Alveodental Consonant [ɬ]? *International Journal of Linguistics and Communications*, 45-65.

Scatton, E. A. (1993). Bulgarian. In B. Comrie, & G. G. Corbett, *The Slavonic Languages* (pp. 188-248). Routledge.