

Nasal place assimilation in Substance Free Logical Phonology

ABSTRACT: This paper provides an analysis of Croatian nasal place assimilation from the perspective of Substance Free Logical Phonology. Treating phonology as application of logico-mathematical functions to abstract, symbolic primitives such as features, the analysis explicitly captures both the fact that in Croatian each underlying nasal behaves differently with respect to place assimilation and the fact that in certain cases nasals assimilate in continuancy as well as in place. On the basis of the provided analysis and on theoretical grounds, it is argued that not all possible *why*-questions that could be asked about phonology are grammatically relevant: True phonological *why*-questions, those that receive explanation within a formal grammar, should be strictly distinguished from functionalist *why*-questions, which concern themselves with aspects of phonology that are reducible to language-external systems and should not be encoded in a generative grammar.

Keywords: generative grammar; substance free logical phonology; nasal place assimilation; continuancy assimilation; formalist vs. functionalist *why*-questions

1. INTRODUCTION

In this paper, I report a previously unrecognized phonological pattern in Croatian, where place assimilation of nasal segments is coupled with continuancy assimilation. I demonstrate that a phonological analysis which adopts a substance-free generative perspective (Hale & Reiss 2008; Reiss 2018) and which treats phonological computation as application of logical functions to unstructured sets of features (Bale et al. 2014; Bale & Reiss 2018), adequately describes the phonological competence of an idealized Croatian speaker with respect to the assimilation of nasals. As I show through appeal to the Degree of Articulatory Constraint (DAC) model of lingual coarticulation (Recasens et al. 1997; Recasens 2018), the characteristic phenomena that functionalist theories of phonology (e.g., Classic Optimality Theory) purport to explain, such as the difference in the propensity of different nasal segments to assimilate (see (1–3) for examples), receive independent and more coherent explanation on purely phonetic grounds.

I argue that phonological *why*-questions of the general format “Why does a

phonological pattern have the form that is does?” are for the better part to be explained on phonetic grounds, and that encoding the answers to such questions in a generative grammar (i.e., claiming that these answers are part of a speaker’s implicit phonological knowledge) can be a theoretical and a methodological mistake.

2. CROATIAN NASAL PLACE ASSIMILATION

Croatian has three underlying nasal segments – /n/, /m/ and /ɲ/ – and all three behave differently with respect to place assimilation. As shown in the following examples, /n/ assimilates to bilabials, labiodentals and velars (1); /m/ assimilates only to labiodentals (2); /ɲ/ does not assimilate (3).¹

(1) /n/-assimilation

- a. /jedan-put/ → [jedamput] ‘once’
- b. /tʃin-b-en-ik/ → [tʃimbenik] ‘factor’
- c. /on prat-i/ → [omprati] ‘he follows’
- d. /invalid/ → [imɲvalid] ‘invalid’
- e. /on vid-i/ → [omɲvidi] ‘he sees’
- f. /kon-form-iz-am/ → [komɲformizam] ‘conformity’

¹ The representative examples from (1) through (3) have been compiled from and verified in the following sources on Croatian: Brabec et al. (1968: §57); Barić et al. (1979: §59–§62, §117–§118); Raguž (1997: §40); Barić et al. (2003: §139–§144); Silić & Pranjković (2005: 25–27); Težak & Babić (2007: §132–§134); Škarić (2007: §246–§247); Težak (2007: §387); Bičanić et al. (2013: 146); Marković (2013: §3).

- g. /bank-a/ → [baŋka] ‘bank’
- h. /kongres/ → [koŋgres] ‘congress’
- i. /inxibir-a-ti/ → [iŋxibirati] ‘to inhibit’

(2) /m/-assimilation

- a. /tramvaj/ → [traŋvaj] ‘tram’
- b. /amfor-a/ → [aŋfora] ‘amphora’
- c. /pitam vas/ → [pitaŋvas] ‘I am asking you’
- d. (/iznim-k-a/ → [iznimka] ‘exception’; *[izniŋka])
- e. (/kamp/ → [kamp] ‘camp’)

(3) absence of /ŋ/-assimilation

- a. (/koŋ bi/ → [koŋbi] ‘horse would’; *[kombi])
- b. (/voŋ te mutf-i/ → [voŋte mutfi] ‘the smell bothers you’; *[vonte mutfi])
- c. (saŋk-e → [saŋke] ‘sled’; *[saŋke])

In underlying representations, the symbol ‘-’ denotes a morpheme boundary within a word, and a blank space denotes a boundary between words. Croatian nasal place assimilation is active across both of these boundaries. In (2d), (2e) and (3a–c) the bracketed examples show the absence of place assimilation in certain contexts: /m/ surfaces unchanged before bilabials and velars, and /ŋ/ surfaces unchanged in all contexts.

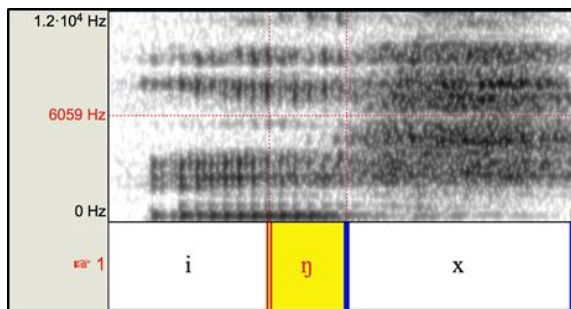
What examples from (1) through (3) do not show, and what is invariably absent from grammatical descriptions of Croatian (e.g., from all sources listed in footnote 1), is the fact that *in some cases* place assimilation is coupled with continuancy assimilation. In other words, a nasal stop sometimes becomes a nasal continuant. This is regularly observed in two contexts: before labiodentals [f] and [v], and before a velar fricative [x].

The Croatian labiodental nasal [m] does not occur in underlying representations; it is a contextually conditioned allophone of

/n/ and /m/. Phonetically, [m] is realized as a continuant, that is, without a complete obstruction of the airflow in the oral cavity. While [m] is sometimes described as a nasal *stop* in phonetic literature (e.g., Laver 1994: 215; Rogers 2000: 194; Ashby & Maidment 2005: 54–55), this continuant realization of [m] is in line with Ladefoged’s & Maddieson’s (1996: 18) careful consideration: “We do not know if a true occlusive could be made with [a labiodental] gesture, when we take into account the gaps that often occur between the incisors.” Apparently, the default articulation of the Croatian [m] entails oral continuancy. Indeed, a Croatian source on phonetics, Škarić (2007: 72), recognizes the fact that Croatian [m] is articulated as a continuant sound. However, this individual observation has not been integrated in the broader explanation of what goes on with Croatian nasals when they assimilate in place.

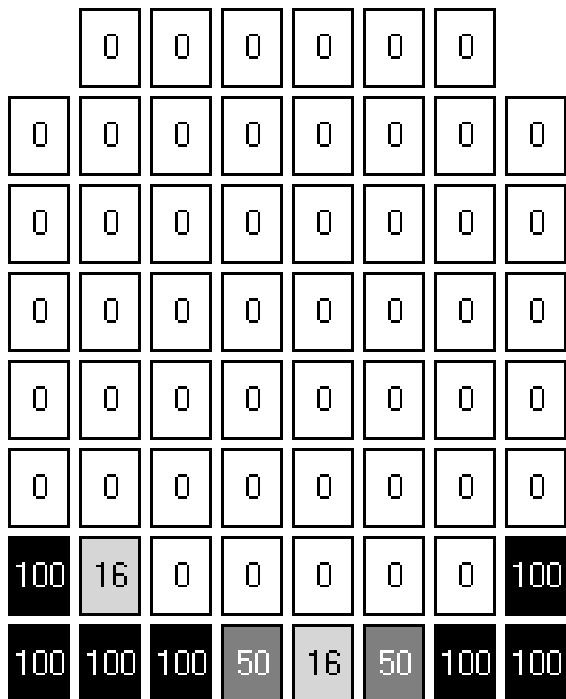
Another context where the nasal changes from a stop to a continuant is before /x/. Thus, while /n/ remains a stop before velars /k/ and /g/, it assimilates in continuancy as well as in place before /x/. An unfortunate state of affairs is that the IPA contains only a single symbol for a velar nasal segment, namely /ŋ/, which is traditionally interpreted as a velar nasal *stop*. An *ad hoc* symbol such as /ŋ̃/ is inappropriate since it incorrectly suggests a fricative articulation of the nasal. Rather, the nasal in a Croatian [ŋx] cluster, such as that in (1i), is realized as an approximant, that is, without frication. The spectrogram in *Figure 1* shows the pronunciation of the beginning of the word [iŋxibirati] (example (1i)) by a native speaker. It can be observed that during the pronunciation of the nasal a clear formant structure without any substantial high-frequency aperiodic sound (corresponding to a fricative) is present. The frication that ensues, corresponding to [x], no longer contains nasality. So the [ŋ] in [ŋx] is not a fricative, but rather an approximant.

Figure 1. A spectrogram corresponding to the pronunciation of the string [iŋx] of the Croatian word *inhibirati* ‘to inhibit’. There is no frication during the nasal part (yellow), indicating that the nasal is an approximant and not a fricative.



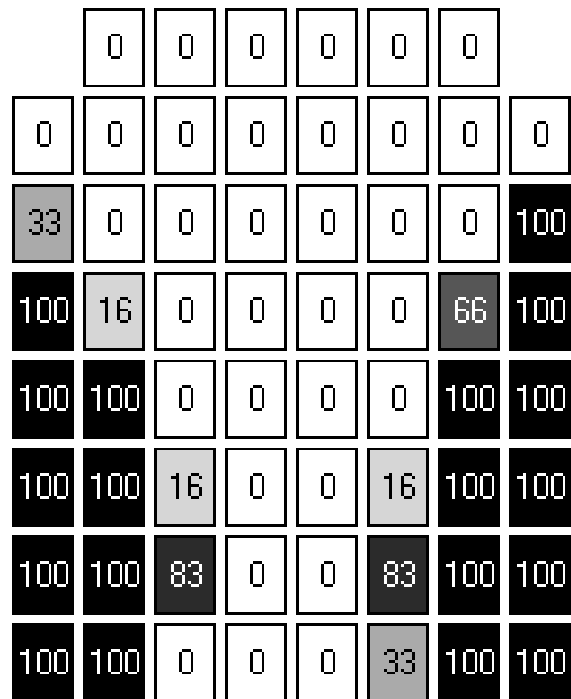
While examples (1g), (1h) and (1i) all contain an assimilated [ŋ], a pertinent phonetic difference is that in (1g) and (1h) [ŋ] is a non-continuant, while in (1i) [ŋ] is a continuant. A preliminary electropalatographic (EPG) investigation

Figure 2. [ŋ] in [baŋka]. The nasal is a non-continuant (i.e., a stop), as indicated by the shading of every cell in the bottom row, which corresponds to a complete oral constriction at the velum.



supports this claim. *Figure 2* shows an EPG record of the pronunciation of the relevant part of the word [baŋka], while *Figure 3* shows an EPG record of the pronunciation of the relevant part of the word [iŋxibirati]. For the purpose of EPG recording, the words were embedded in sentences in order to remove the focus from what is being studied and thus to elicit a spontaneous pronunciation; each sentence was pronounced six times by the same native speaker, and the results were then averaged for each of the two words. In electropalatograms in *Figures 2* and *3*, white cells denote the absence of tongue-to-palate contact; shaded cells denote tongue-to-palate contact.² The EPG frame presented in both *Figures* is the frame approximately corresponding to the middle of the duration of the nasal.

Figure 3. [ŋ] in [iŋxibirati]. The nasal is a continuant, as indicated by the absence of shading in the middle two columns, which corresponds to unobstructed airflow in the medial region of the oral cavity.



² When EPG data is averaged across multiple repetitions of an utterance by the same speaker, as is the case in *Figures 2* and *3*, the numbers in the shaded cells tell us the percentage in which each electrode of the artificial palate was activated by the tongue.

Figure 2 is an EPG record of the pronunciation of the medial part of the nasal in the [ŋk] cluster of the word *banka*. Shaded squares depict complete closure in the oral cavity at the velum, confirming that the nasal is a non-continuant before the velar stop [k].³ Figure 3 is an EPG record of the pronunciation of the medial part of the nasal in the [ŋx] cluster of the word *inhibirati*. It is clear that while there is a *narrowing* at the velum (particularly noticeable in the next to last row in Fig. 3), there is a consistent absence of a complete closure in the oral cavity (as indicated by two middle columns containing only white cells), confirming that the nasal is a continuant before [x]. Figures 1 through 3 confirm that the Croatian [ŋ] is a velar continuant approximant in a [ŋx] cluster.

To summarize, each of the three Croatian nasal phonemes behaves differently with respect to place assimilation ((1) – (3)), and in specific cases ((1d), (1e), (1f), (1i), (2a), (2b), (2c)) Croatian nasals assimilate in continuancy as well as in place.

3. CROATIAN NASAL PLACE ASSIMILATION IN SUBSTANCE FREE LOGICAL PHONOLOGY

The main tenet of Substance Free Logical Phonology (SFLP) is that phonological computation proceeds irrespective of phonetic substance; rather, phonological computation is the application of logical functions to sets of abstract symbols, such as features and other phonological primitives (Hale & Reiss 2008; Bale et al. 2014; Reiss 2018; Bale & Reiss 2018). This *does not* translate to the claim that the relation between features and their phonetic correlates is arbitrary; it just means that this lawful relation is phonologically irrelevant (see Volenec & Reiss 2017).

In SFLP, *segments* are sets of valued features, and *natural classes* are sets of sets of valued features. A *set* is a mathematical notion—a well-defined collection of distinct members. One important property of a set is that it is an unstructured, unordered collection of members. So the set {+SONORANT, +NASAL} is equal to the set {+NASAL, +SONORANT}. Slightly departing from the traditional practice, henceforth I will enclose sets of valued features (i.e., segments) in curly brackets, and sets of sets of features (i.e., natural classes) in square brackets. A natural class of segments can be defined through a minimal set of features that are shared by all segments in that class and no other segments. For example, [+NASAL] will refer to all and only those segments in a given language which contain the feature +NASAL; in Croatian, [+NASAL] = { {n}, {m}, {ŋ}, {mŋ}, {ŋ} }.

Phonological computation works by way of logical functions. A *function* is a relation that associates each member x of a set X (the *domain* of the function) to a single member y of another set Y (the *codomain* of the function, which may be identical to the domain). In phonology, the domain and the codomain are strings of segments. Two functions that I will make use of in this paper are *subtraction* and *unification*, defined in (4) and (5), respectively.

(4) Subtraction

If A and B are sets, then $A - B$ results in the set that contains all and only the members of A that are not members of B .

(5) Unification

If A and B are sets, then $A \sqcup B$ results in the smallest set that contains all the members of A and all the members of B . In order for $A \sqcup B$ to be defined, the result must be consistent.

As can be seen from the definition in (5), the result of unification meets the requirement

³ On a side note, the 16 % cell in the bottom row of Figure 2 does not necessarily mean that [ŋ] in a [ŋk] cluster is sometimes realized as a continuant; it is more likely that the complete obstruction consistently occurred deeper in the oral cavity, beyond the reach of the artificial palate.

of *consistency*, which is defined in (6) (Bale & Reiss 2018: 377).

(6) Consistency

A set of features ρ is consistent if and only if there is no feature F such that $+F \in \rho$ and $-F \in \rho$.

As defined in (6), consistency merely ensures that unification does not yield logically incompatible feature sets, such as $\{+VOICED, -VOICED\}$. Note that consistency ensures the absence of *logically* incompatible feature values, and otherwise says nothing about the incompatibility in valued feature combinations. A segment that is, say, $\{+NASAL, -SONORANT\}$ is not excluded on *phonological* grounds, that is, such a segment is a phonologically (i.e., cognitively) well-defined (i.e., possible) segment. Rather, such a segment is excluded on *phonetic* grounds, since it imposes impossible demands upon the articulatory apparatus and is therefore phonetically uninterpretable. Since it is phonetically uninterpretable, it is unpronounceable; since it is unpronounceable, it will never occur in the primary linguistic data; since it never appears in the primary linguistic data, it will never be acquired by a language learner; since it will never be acquired (and is not innate), it will never be part of any I-language. While the absence of a $\{+NASAL, -SONORANT\}$ segment from an I-language *is* something to be accounted for in a generative grammar (by not including it in an I-language’s phoneme inventory), *the reason* for the absence of such a segment is not, since the reason is phonetic, that is, extra-grammatical. On the other hand, both the absence of a $\{+VOICED, -VOICED\}$ segment and the reason for its absence is to be accounted for in a generative grammar, since the reason is consistency, a characteristic of how unification works, which is part of grammar.

With these preliminaries in mind, we can now postulate the SFLP functions that account for nasal place assimilation in

Croatian. It is worth emphasizing that in generative phonology the object of study is an aspect of the mind/brain which, in interaction with other systems, yields observable data (such as those in (1) – (3)); data itself is not the object of study, data is evidence for inferring about the actual object. Therefore, the *general* aim of a phonological analysis is not to ‘account for the data’, but rather to provide an explicit characterization of the relevant cognitive fragment, using data as evidence. The *specific* aim of the phonological analysis in this paper is to provide an explicit characterization of the cognitive fragment that governs the assimilation of nasals, using data in (1) – (3) as evidence.

Treating assimilation as a two-step process (Harris 1984; Samuels 2011; Bale et al. 2014), the functions in (7a) and (7b) subtract relevant sets of features from n and m ; by means of set unification, (8a) then captures the assimilation to bilabials and labiodentals, while (8b) captures the assimilation to velars.

(7a) $[+NAS, +ANT] - \{+COR, -LAB, -CONT\} / _ [+LAB]$

(7b) $\{+NAS, +ANT, +COR\dots\} - \{+ANT, +COR, -BACK, -CONT\} / _ [-SON, +BACK]$

(8a) $[+NAS, +ANT] \sqcup \{-COR, +LAB, \alpha CONT\} / _ [+LAB, \alpha CONT]$

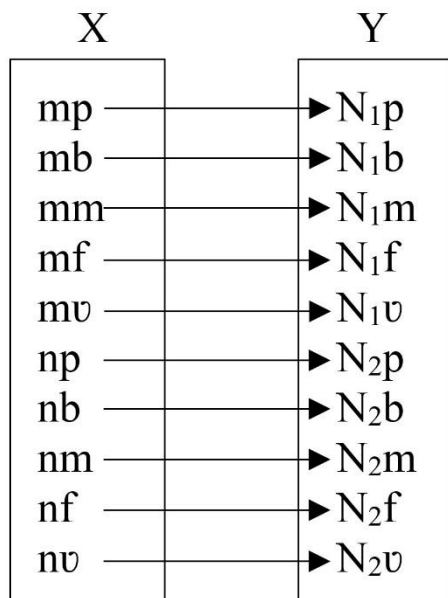
(8b) $[+NAS] \sqcup \{-ANT, -COR, +BACK, \alpha CONT\} / _ [-SON, +BACK, \alpha CONT]$

Function (7a) targets the natural class $[+NAS, +ANT]$, which is a set that contains the members $\{m\}$, $\{n\}$, and $\{ŋ\}$, each of which is a set of valued features. The environment in which (7a) applies is defined by the natural class $[+LAB] = \{\{p\}, \{b\}, \{m\}, \{ŋ\}, \{f\}, \{v\}, \{v\}\}$. In Croatian, $\{ŋ\}$ occurs neither in underlying representations nor in any derived representation where it could serve as a target of another rule (Volenc 2018), so $\{ŋ\}$ will never be targeted by (7a). Similarly, $\{v\}$ (not to be confused with $\{v\}$)

does not occur in Croatian underlying representations but only as a voiced allophone of {f}, and will thus never actually participate in (7a). The function (7a) subtracts the features +COR, –LAB and –CONT from members of the target, yielding an incomplete (underspecified) nasal segment. Note that when the target is {m}, (7a) only subtracts the feature –CONT, since {m}, unlike {n}, contains neither +COR nor –LAB. Thus the output of function (7a) will vary depending on which nasal segment is the target. If {m} is the target, then the underspecified nasal segment will contain –COR and +LAB; I will refer to this segment as N_1 . If the target is {n}, then the output will be unspecified for both COR and LAB; I will refer to this segment as N_2 .

Like any function, (7a) has a domain, call it X , and a codomain, call it Y ; (7a) maps members of X to members of Y in a particular way, which can be represented as in *Figure 4*.

Figure 4. A mapping diagram of the function (7a).

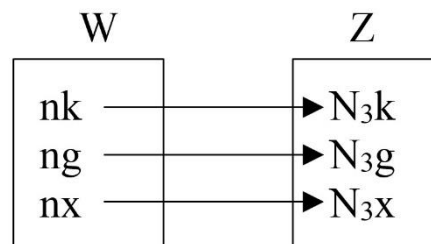


In brief, (7a) can informally be interpreted as ‘remove the features +COR and –CONT from any member of the set {{m}, {n}, {ŋ}} if that member is in the context before a member of the set {{p}, {b}, {m}, {f}, {v}, {v}}’. *Figure 4* reflects

the fact that not all combinations/mappings are attestable in the data due to the fact that {ŋ} and {v} are not phonemes in Croatian.

Function (7b) does not target a natural class (a set of sets) but rather a single set of features, namely the set {+NAS, +ANT, +COR...}, which corresponds to the segment {n}. The three features listed are sufficient to characterize the set (hence the ‘...’), but it should be kept in mind that {n} corresponds to a full set of valued features (see the table in (9)). Function (7b) targets only {n} because in Croatian only that segment assimilates to velars. The environment in which (7b) applies is defined by the natural class [–SON, +BACK] = {{k}, {g}, {x}}, that is, it applies before velar obstruents. The function (7a) subtracts the features +ANT, +COR, –BACK and –CONT from {n}, yielding an underspecified nasal segment, which I will refer to as N_3 , since it is different from both N_1 and N_2 . In this case, the domain (W) of the function consists of strings nk , ng , nx , while the codomain (Z) consists of strings N_3k , N_3g , N_3x . The mapping in question is represented by the diagram in *Figure 5*.

Figure 5. A mapping diagram of the function (7b).



In brief, (7b) can informally be interpreted as ‘remove the features +ANT, +COR, –BACK and –CONT from {n} if that set is in the context before a member of the set {{k}, {g}, {x}}’. In order to facilitate keeping track of the nasal segments involved in these operations, the table in (9) provides feature specifications of the nasals that I have thus far referred to.

(9) Feature specifications for nasals

	m	ɱ	n	N ₁	N ₂	N ₃
CONS	+	+	+	+	+	+
SON	+	+	+	+	+	+
COR	-	-	+	-		
ANT	+	+	+	+	+	
LAB	+	+	-	+		-
NAS	+	+	+	+	+	+
CONT	-	+	-			
BACK	-	-	-	-	-	

Function (8a) targets those anterior nasals that occur before labial segments. Since function (8a) is ordered *after* function (7a), all anterior nasals in that position have already been subjected to subtractions of (7a), and (8a) will therefore apply to strings containing underspecified segments N_1 and N_2 . In other words, the domain of (8a) is identical in members to the codomain Y of the function (7a) (see Fig. 4). (8a) performs the unification of N_1 and N_2 with the set of features $\{-COR, +LAB, \alpha CONT\}$. The value of the variable α assigned to CONT will be determined by the value that appears in the context. There are, of course, two possibilities.

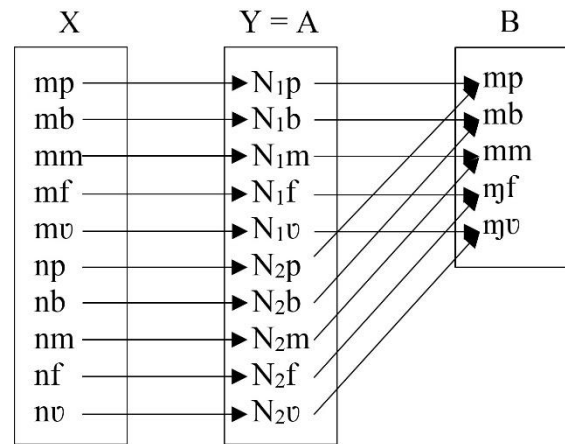
If $-CONT$ is in the context, then the function will unify N_1 and N_2 with $-CONT$. In other words, before $\{p\}$, $\{b\}$, $\{m\}$ the nasals will map to the bilabial stop $\{m\}$. N_1 is already $-COR, +LAB$, so it will just adopt $-CONT$; N_2 will adopt all three features.

If $+CONT$ is in the context, then the function will unify N_1 and N_2 with $+CONT$. In other words, before $\{f\}$, $\{v\}$ the nasals will map to the labiodental continuant $\{ɱ\}$. Note that the feature that distinguishes between bilabial and labiodental segments is the feature CONT. In general, the feature STRIDENT can distinguish between bilabial fricatives $[\phi, \beta]$ (which are $-STRID$) and labiodental fricatives $[f, v]$ (which are $+STRID$). However, the feature STRIDENT is not appropriate for distinguishing between Croatian bilabials and labiodentals, since the class of Croatian labiodentals contains the sonorant $[v]$, which is $-STRID$ (like all

Croatian bilabials). So the only feature that consistently distinguishes between all Croatian bilabial and labiodental segments is the feature CONT.

Functions (7a) and (8a) are connected to the extent that they share a domain: the codomain Y of (7a) is the same as the domain A of (8a). (8a) then maps members of the domain A to members of the codomain B . These relationships are represented in Figure 6.

Figure 6. A mapping diagram of functions (7a) and (8a).

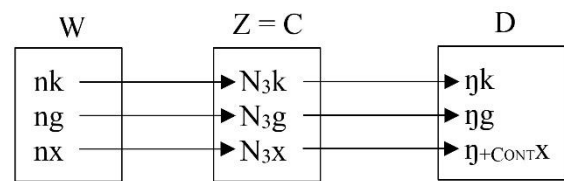


Function (8b) targets nasal segments in the context before velar obstruents, and unifies them with the set of features $\{-ANT, -COR, +BACK, \alpha CONT\}$. The value of the feature CONT is determined by the context (see below). Note that only N_3 will actually successfully undergo unification. N_1 and N_2 will never occur in this context (since they occur only in representations derived by (7a)), and neither will $\{ɱ\}$ (since it is not a phoneme). Because (7b) is ordered before (8b), $\{n\}$ will always first be subjected to subtraction in this context, so it also cannot serve as the input to (8b). The remaining two nasal phonemes in Croatian, $\{m\}$ and $\{ɲ\}$, will fail to unify due to violating consistency (6). For example, $\{m\}$ contains the members $+ANT$ and $-BACK$, and (8b) requires unification with $-ANT$ and $+BACK$. According to consistency, the same segment cannot be both $+ANT$ and $-ANT$, therefore the function will be undefined in

such cases and will not yield an output. So (8b) predicts that underlying /mk/, /mg/, /mx/, /ɲk/, /ɲg/, /ɲx/ will surface as [mk], [mg], [mx], [ɲk], [ɲg], [ɲx], respectively, which is exactly what the evidence in (2) and (3) suggests.

N_3 is unspecified for features ANT, COR, BACK and CONT (see the table in (9)), and will therefore not violate consistency when function (8b) is applied. Through (8b), N_3 will adopt the valued features –ANT, –COR, +BACK. This amounts to *place* assimilation, as N_3 changes from a placeless nasal to a velar nasal before velar obstruents. Furthermore, (8b) predicts that N_3 will also assimilate in *continuancy*. If the natural class in the context of (8b) is [–SON, +BACK, –CONT], that is, if it contains the members {k} and {g}, then N_3 will adopt –CONT via the α variable. In this case, the output will be a velar nasal *stop*. However, if the natural class in the context of (8b) is [–SON, +BACK, +CONT], that is, if it contains the member {x}, then N_3 will adopt +CONT via the α variable. Here, the output will be a velar nasal *continuant*. Since the IPA does not have distinct symbols for these two segments, I will use the traditional [ɲ] for the velar nasal stop, and the *ad hoc* [ɲ+CONT] for the velar nasal continuant.

Figure 7. A mapping diagram of functions (7b) and (8b).



Functions (7b) and (8b) are connected to the extent that they share a domain: the codomain Z of (7b) is the same as the domain C of (8b). (8b) then maps members of the domain C to members of the codomain D . These relationships are represented in *Figure 7*.

Functions (7a), (7b), (8a), (8b) explicitly model phonological competence of an idealized Croatian speaker with respect to the assimilation of nasals. It is worth reiterating that (7a) is ordered before (8a), and (7b) is ordered before (8b); other relationships, such as that between (7a) and (7b), are undetermined because those functions are independent of each other. As the derivation diagram in *Table 1* shows, the predictions that these ordered functions give are in line with the data/evidence in (1) – (3), including the absence of assimilation in particular cases as well as the coupling of continuancy assimilation with place assimilation.

Table 1. Derivations of relevant Croatian forms.

URs	banka	inxibirati	on bi	invalid	tramvaj	iznimka	kamp	saŋke	koŋ bi
(7a)	–	–	oN ₂ bi	iN ₂ valid	traN ₁ vaj	–	kaN ₁ p	–	–
(7b)	baN ₃ ka	iN ₃ xibirati	–	–	–	–	–	–	–
(8a)	–	–	ombi	iɲvalid	traɲvaj	–	kamp	–	–
(8b)	baɲka	iɲ _[+CONT] xibirati	–	–	–	–	–	–	–
SRs	baɲka	iɲ _[+CONT] xibirati	ombi	iɲvalid	traɲvaj	iznimka	kamp	saŋke	koŋbi
Gloss	‘bank’	‘to inhibit’	‘he would’	‘invalid’	‘tram’	‘exception’	‘camp’	‘sled’	‘horse would’

4. DISCUSSION: EXPLANATION IN PHONOLOGY

Functionalist theories of phonology, such as Classic Optimality Theory, assume that it is in their purview to explain why a particular phonological pattern that exists in the mind of a speaker has the form that it does (McCarthy 2002: §4.4). For example, inspecting the data in (1) – (3), it is clear that the coronal nasal assimilates in more contexts than the bilabial nasal, which in turn assimilates in more contexts than the inert palatal nasal. So with respect to the propensity of Croatian nasals to assimilate, the following hierarchy, which is commonplace in phonological typology, can be constructed: coronal nasal > bilabial nasal > palatal nasal. Here the symbol '>' merely means 'assimilates in more contexts than'.

However, including the answers to these kinds of *why*-questions in the phonological grammar is a mistake for at least two reasons. First, it is implausible to claim that the *implicit* phonological knowledge of a speaker contains an *explicit* explanation of the reasons for which the pattern has the form it has, just as, say, a Croatian speaker's morphosyntactic knowledge does not contain an explanation of *why* the dative of the Croatian word *stol* 'table' is the form *stolu* 'to the table' and not any of the infinitely many other possible forms. From the point of view of phonological competence, the aforementioned hierarchy is irrelevant—it is an intellectual creation of the linguist and not of the language learner—and it would thus be a mistake to attribute it to the phonological module of the grammar.

It has notably been stated that separating the causes of phonological processes from the mechanisms that yield them, that is, separating the *why* from the *what*, "will mark a major defeat for the [phonological] enterprise" (Prince & Smolensky 1993/2004: 234). McCarthy (2002: 221) added that "the problem with

this [formal, rule-based—*v*] approach is that linguistic theory – the formal side of things – ends up doing very little of the explaining, and the external postulates end up doing almost all of it." But such a conception of the generative phonological enterprise disregards the fact that "where properties of language can be explained on functional grounds, they provide no revealing insight into the nature of the mind." (Chomsky 1971: 44).

Indeed, phonological *why*-questions can be divided into two categories, which I will refer to as *formalist why*-questions and *functionalist why*-questions. The first category is concerned with phonological UG, that is, with the general question 'Why does phonology have the computational properties that it does?'⁴ The second category is concerned with how phonology interacts with other systems (e.g., the perceptual system, the motor system etc.) in language use, which then leads to observable data. The crucial distinction is that the answers to the first type of questions can be conceived as properties of the mental grammar itself, while the answers to the second questions are properties not of the mental grammar but rather of the systems with which the grammar interacts. Obviously, for the study of *language*—as distinct from *language use*—the formalist *why*-questions have epistemological priority; also, the answers to those questions *should* be encoded in a generative grammar because that information characterizes linguistic knowledge. A concrete example of answering a formalist *why*-question within the grammar can be drawn from closely inspecting function (8b), which accounts for the $\{N_3\} \rightarrow \{\eta\}$ mapping. Consider the question 'Why is the target of function (8b) defined so generally, as the natural class [+NAS], if it ignores every other nasal segment and applies only to N_3 ?' This question is just a specific version of the question 'Why does function (8b) have the form that it does?'. The answer to this

⁴ Of course, along with the general *what*-question, namely 'What are the computational properties of phonology?'

question is a formal property of grammar: Phonological functions target natural classes, and an underspecified segment cannot define a natural class alone (see Bale & Reiss 2018 for details, particularly §45.2 and §52).⁵ Since N_3 is underspecified, it cannot be the sole target of a phonological function. If correct, this universal principle, then, is truly revealing of how the mind works since it cannot be reduced to the properties of the sensorimotor system. To take another example, the reason for why phonological processes never generate segments that contain {+VOICED, -VOICED} are to be found in the formal properties of functions, that is, they come ‘from within’ and not from “external postulates”. Consistency, as defined in (6), is just a formal property of how set unification works.

It must be emphasized that not admitting functional considerations into grammar does not mean that “external postulates”, as McCarthy (2002: 221) calls them, cannot or should not serve as evidence for inferring about the nature of grammar. In principle, linguistic theory can draw evidence about its object of study from wherever—there is no *a priori* way to determine what may count as evidence in explaining the human language faculty and there is no reason to discard a piece of evidence simply because of its source. If physicists can legitimately draw evidence about supernovae from studying ceramic rabbit-pots (Antony 2003: 55–58), then surely it is not a “problem”, as McCarthy (2002: 221) suggests, that linguistics draws evidence from phonetics while not admitting phonetic substance into the theory. The relative proportion of conclusions drawn from “external postulates” as opposed from within linguistic theory (i.e., from formal properties of grammars) is not only something that cannot be measured in any coherent way but it is also completely irrelevant. If it turns out that 99 % of

properties of all I-languages can be explained by “external postulates”, then so be it—that does not constitute a scientific problem in itself. However, judging by the repeated failure to reduce phonology to functional phonetics (Sapir 1933/1949; Hale & Reiss 2008; Samuels 2011), and also by the explanatory success achieved on purely formal grounds in other linguistic domains (Chomsky 1995), it is much more likely that formal properties of grammars do play a significant role in the shaping of an I-language. The point is that while something may serve as evidence about our object of study, we must not automatically encode the evidence itself into our explanatory theory. The fact that ceramic rabbit-pots serve as astronomical evidence does not mean that theories of supernovae should contain reference to rabbit-pots. Likewise, the fact that speech production and perception serve as linguistic evidence does not mean that theories of language should include reference to speech production and perception. If we are committed to the study of phonology as a branch cognitive science, then encoding the answers to the functionalist *why*-questions in a phonological grammar is a theoretical error.

The second reason not to encode the answers to the functionalist *why*-questions in a phonological grammar is of a more general scientific nature: It serves no actual purpose as it does not advance our understanding of the phenomenon in question. To return to a concrete example—the question of why the Croatian coronal nasal assimilates in more contexts than the bilabial nasal, which in turn assimilates in more contexts than the palatal nasal (1–3)—an explanation for such a pattern is readily available in phonetics. The core principle of the Degree of Articulatory Constraint model of lingual coarticulation is the following: Coarticulatory resistance and coarticulatory influence of a speech sound rise in proportion to the degree of the tongue back involvement in articulating a given

⁵ Note that while (7b) targets a single segment, that function is still in line with this principle since a single fully specified segment can always define a natural class.

sound (Recasens et al. 1997; Recasens 2018: §2.7). Articulating the palatal nasal [ɲ] significantly involves the biomechanically inert tongue back, therefore [ɲ] will be particularly resistant to coarticulatory adaptations to adjacent sounds. On the other hand, articulating the coronal nasal [ɲ] involves the less constrained tongue front, therefore [ɲ] will be particularly prone to coarticulatory adaptations to adjacent sounds. It follows that the less inert [ɲ] will change in more contexts than the more inert [ɲ]. Also, since the coronal nasal shares its active articulator (i.e., the tongue) with the velars, [ɲ] adapts to the velars, and this coarticulatory effect is apparently phonologized as part of the function in (8b), while [m], which employs the lips as the active articulator, is independent from the velars and does not adapt to them. Thus phonetics provides an answer to the functionalist *why*-question. It should be emphasized that there are phonological patterns which defy systematic phonetic explanation, leading to phonetic unnaturalness: For example, the coronal nasal does not share the active articulator with the labials and yet in Croatian it adapts to them (see 1a–f). In that particular case, it is clear that phonetics (i.e., the movement of the articulators) is grounded in phonology (i.e., driven by mental operations (7a) and (8a)) and not *vice versa*. Since the explanation for such patterns is for the better part already provided by phonetic research, importing this insight into a generative grammar (for example, in the form of a markedness constraint) adds nothing new to our understanding. Furthermore, grounding

these imported explanatory devices in typology in order to augment either their plausibility, their explanatory breadth, or both,⁶ fares no better. Consider the logic of postulating a specific, well-known markedness constraint. One starts by observing that in many languages syllables do not have codas.⁷ One can now ask *why* that is the case; in other words, this observation demands an explanation. The observation about this tendency is converted into a markedness constraint NO-CODA, defined as ‘assign one violation mark for every segment in a syllable coda’.⁸ In languages where this tendency is overtly manifested in surface representations, such as Hawaiian where there are only V and CV syllables, NO-CODA is ranked sufficiently highly and the grammar generates only output forms without codas. Thus OT provides an answer to the question why in many languages syllables do not have codas: because in these languages NO-CODA has a high rank, that is, it is not dominated by other syllable structure constraints. But this is not an *explanation* of the initial observation, it is merely a restating of the observation in a different format (in the format of a ranked constraint). In other words, that which demands an explanation is being explained by referring to a slightly mutated version of itself. An ‘explanation’ where the *explanandum* and the *explanans* are the same is circular, and therefore worthless. With respect to NO-CODA, the *explanandum*, i.e., that which demands an explanation, is the observation that in many languages syllables do not have codas. The *explanans*, i.e., that which serves as an explanation, is the sufficiently high ranking

⁶ „Since OT is a theory of grammar, the consequences are displayed in the grammars predicted and disallowed – ‘typological evidence’. *A constraint which cannot be justified on those grounds cannot be justified.*” (Prince 2007: 46; emphasis added—vv)

⁷ „The poster children of typologically well-grounded constraints are Onset and No-Coda. Many languages don’t allow syllables to start in a vowel and many languages don’t display syllables closed by a consonant.” (Krämer 2018: 39)

⁸ Here is an example of such reasoning from an influential OT textbook (Kager 1999: 94):

Both language typology and the wide-spread occurrence of processes which avoid codas suggest that the ‘unmarked’ situation is for syllables to lack codas. This unmarked situation is encoded in the following well-formedness constraint.

(7) No-Coda (‘Syllables are open.’)

of NO-CODA. But here the *explanans* does not explain anything, it merely forces us to reformulate the original question: Why, then, is NO-CODA ranked highly in many languages? To truly answer the question, we have to look elsewhere (e.g., in how speech perception works), which is exactly what we would have to do anyway (i.e., if we maintained a completely formal linguistic theory) *if* for some reason we were determined to answer these kinds of functionalist *why*-questions. Far-reaching typological implications are widely considered to be a defining strength of OT:

OT is inherently typological: the grammar of one language inevitably incorporates claims about the grammars of all languages. This joining of the individual and the universal, which OT accomplishes through ranking permutation, is probably the most important insight of the theory. (McCarthy 2002: 1)

[C]onstraint interaction is the source of cross-linguistic variation. (Krämer 2018: 41)

But OT has no *actual* implications for typology: What interacts is an inordinate number of observed negatively-formulated tendencies (i.e., violable constraints), each of which either demands or already has an *independent explanation*. Crucially, every typologically grounded constraint has been defined on the basis of observing cross-linguistic variation (otherwise it would not be typologically grounded). So “the most important insight of [optimality] theory” (McCarthy 2002: 1) is that the “source of cross-linguistic variation” (Krämer 2018: 41) is the interaction of observations about cross-linguistic variation. Such circular reasoning does not explain anything and is bare of any scientific value; even worse, it perpetually obfuscates the issue by confusing scholars into believing that the explanation has been attained. Put briefly, “the only way to keep a reasoning non-circular is to make sure that the data on which it has been built and those on which it makes a prediction are independent.” (Scheer 2004: 475).

5. CONCLUSION

The proposed SFLP analysis in (7) and (8) explicitly describes the phonological competence of an idealized Croatian speaker with respect to the assimilation of nasals. In particular, the formal analysis provided in this paper captures the fact that each of the three Croatian nasal phonemes behaves differently with respect to place assimilation ((1) – (3)), and that in specific cases ((1d), (1e), (1f), (1i), (2a), (2b), (2c)) Croatian nasals assimilate in continuancy as well as in place.

The explanation for the difference in the propensity of the nasals to assimilate is provided solely on phonetic grounds, through appeal to the DAC model of lingual coarticulation. Restating this phonetic explanation in the phonological module of the grammar (e.g., through the vague notion of markedness) without adding something new to it, is therefore an unmotivated violation of the principle of scientific simplicity. It should be noted, however, that there are aspects of the phonological pattern in (1) – (3) that are phonetically unnatural. For example, while [m] does not assimilate to segments containing +CORONAL, [n] does assimilate to segments containing +LABIAL, without any apparent phonetic motivation.

Finally, I have argued that true phonological *why*-questions, those that receive explanation within a formal grammar, should be strictly distinguished from functionalist *why*-questions, which concern themselves with aspects of phonology that are grammatically irrelevant. Keeping these two domains separate leads to simpler and sharper theories about the nature of the language faculty.

6. REFERENCES

Antony, Louise M. 2003. Rabbit-pots and supernovas: On the relevance of psychological data to linguistic theory. In Barber, A. (ed.), *Epistemology of*

- Language*. 47–68. Oxford: Oxford University Press.
- Ashby, Michael, John Maidment. 2005. *Introducing Phonetic Science*. Cambridge: Cambridge University Press.
- Bale, Alan, Maxime Papillon & Charles Reiss (2014) Targeting Underspecified Segments: A Formal Analysis of Feature-changing and Feature-filling Rules. *Lingua* 148: 240–253.
- Bale, Alan & Charles Reiss. 2018. *Phonology. A Formal Introduction*. Cambridge: MIT Press.
- Barić, Eugenija, Mijo Lončarić, Dragica Malić, Slavko Pavešić, Mirko Peti, Vesna Zečević, Marija Znika. 1979. *Priručna gramatika hrvatskoga književnog jezika*. Zagreb: Školska knjiga.
- Barić, Eugenija, Mijo Lončarić, Dragica Malić, Slavko Pavešić, Mirko Peti, Vesna Zečević, Marija Znika. 2003. *Hrvatska gramatika*. Zagreb: Školska knjiga.
- Bičanić, Ante, Anđela Frančić, Lana Hudeček & Milica Mihaljević. 2013. *Pregled povijesti, gramatike i pravopisa hrvatskoga jezika*. Zagreb: Croatica.
- Brabec, Ivan, Mate Hraste & Sreten Živković. 1968. *Gramatika hrvatskosrpskoga jezika*. Zagreb: Školska knjiga.
- Chomsky, Noam. 1971. Deep structure, surface structure and semantic interpretation. In Steinberg, D. & L. Jakobovits (eds.), *Semantics: An Interdisciplinary Reader in Linguistics, Philosophy and Psychology*. 183–216. Cambridge: Cambridge University Press.
- Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge: MIT Press.
- Hale, Mark & Charles Reiss. 2008. *The Phonological Enterprise*. Oxford: Oxford University Press.
- Harris, James W. 1984. Autosegmental phonology, lexical phonology and Spanish nasals. In Mark Aronoff & Richard T. Oehrle (eds.), *Language sound structure*. 67–82. Cambridge: MIT Press.
- Krämer, Martin. 2018. Current Issues and Directions in Optimality Theory. Constraints and Their Interactions. In Hannahs, Stephen J. & Anna R. K. Bosch (eds.), *The Routledge Handbook of Phonological Theory*. 37–67. London – New York: Routledge.
- Ladefoged, Peter & Ian Maddieson. 1996. *The Sounds of the World's Languages*. Oxford: Blackwell.
- Laver, John. 1994. *Principles of Phonetics*. Cambridge: Cambridge University Press.
- Marković, Ivan. 2013. *Hrvatska morfonologija*. Zagreb: Disput.
- McCarthy, John J. 2002. *A Thematic Guide to Optimality Theory*. Cambridge: Cambridge University Press.
- Prince, Alan. 2007. The Pursuit of Theory. In Lacy, Paul de (ed.), *The Cambridge Handbook of Phonology*. 33–60. Cambridge: Cambridge University Press.
- Prince, Alan & Paul Smolensky. 1993/2004. *Optimality Theory: Constraint Interaction in Generative Grammar*. Oxford: Blackwell.
- Raguž, Dragutin. 1997. *Praktična hrvatska gramatika*. Zagreb: Medicinska naklada.
- Recasens, Daniel, Maria Pallarès & Jordi Fontdevila. 1997. A model of lingual coarticulation based on articulatory constraints. *Journal of the Acoustical Society of America* 102: 544–561.
- Recasens, Daniel. 2018. *The Production of Consonant Clusters. Implications for Phonology and Sound Change*. The Hague: Mouton.
- Reiss, Charles. 2018. Substance Free Phonology. In Stephen J. Hannahs & Anna R. K. Bosch (eds.), *The Routledge Handbook of Phonological Theory*. 425–452. New York: Routledge.
- Rogers, Henry. 2000. *The Sounds of Language. An Introduction to Phonetics*. London: Routledge.
- Sapir, Edward. 1933/1949. The Psychological Reality of the Phoneme.

- In Mandelbaum, David (ed.), *Selected Writings of Edward Sapir in Language, Culture and Personality*. 46–60. Los Angeles: University of California Press.
- Samuels, Bridget. 2011. *Phonological Architecture: A Biolinguistic Perspective*. Oxford: Oxford University Press.
- Scheer, Tobias. 2004. *A Lateral Theory of Phonology. What Is CVCV, and Why Should It Be?* Berlin – New York: Mouton de Gruyter.
- Silić, Josip & Ivo Pranjković. 2005. *Gramatika hrvatskoga jezika za gimnazije i visoka učilišta*. Zagreb: Školska knjiga.
- Škarić, Ivo. 2007. Fonetika hrvatskoga književnoga jezika. In Katičić, Radoslav (ed.), *Glasovi i oblici hrvatskoga književnoga jezika*. 15–157. Zagreb: Nakladni zavod Globus.
- Težak, Stjepko & Stjepan Babić. 2007. *Gramatika hrvatskoga jezika. Priručnik za osnovno jezično obrazovanje*. 16. izdanje. Zagreb: Školska knjiga.
- Težak, Stjepko. 2007. Morfonologija. In Katičić, Radoslav (ed.), *Glasovi i oblici hrvatskoga književnoga jezika*. 261–276. Zagreb: Nakladni zavod Globus.
- Volnec, Veno. 2018. *Croatian Phonological Alternations in Optimality Theory*. Doctoral dissertation. University of Zagreb.
- Volnec, Veno & Charles Reiss. 2017. Cognitive Phonetics: The Transduction of Distinctive Features at the Phonology-Phonetics Interface. *Biolinguistics* 11: 251–294.