VENO VOLENEC

THE NEUROBIOLOGICAL REALITY OF PHONOLOGICAL FEATURES

ORIGINAL RESEARCH PAPER
https://doi.org/10.17234/9789533791821.16

Distinctive features are a hallmark of modern phonology, a core theoretical construct present in linguistic theories as diverse as functionalist structuralism and formal generative phonology. However, despite decades of sustained research, it is still not clear what features are, where they reside, and how they relate to phonetic substance. In this paper, I propose features be conceptualized as neural circuits with three properties: distinguishability, combinability and efficacy. The function of these circuits is to yield mental representations of phonologically relevant aspects of speech. I also outline a neurolinguistic model of the phonology–phonetics interface called Cognitive Phonetics (CP), which describes how features relate to phonetic substance during speech production. CP proposes that each feature that comprises a surface phonological representation activates speech movements via two simple neural procedures. I show how CP can account not only for the traditionally recognized *inter*segmental coarticulation, but also for the previously less explored *intra*segmental coarticulation, suggesting that the basic units of speech production are transduced features.

Keywords: phonological features, neuroscience, Cognitive Phonetics, phonology–phonetics interface, coarticulation

---

INTRODUCTION

Distinctive feature theory was initially outlined by Roman Jakobson in a lecture delivered in 1928 (summarized in Jakobson, 1962: 3–6) and in an often-overlooked paper from the late 1930s (Jakobson, 1939), and was subsequently elaborated by Jakobson, Fant & Halle (1952) and Jakobson & Halle (1956). Features have played a central role in frameworks as diverse as functionalist structuralism (where they originated from), formal generative phonology (Chomsky & Halle, 1968; Volenec & Reiss, 2020), functionalist generative phonology (Prince & Smolensky, 1993; McCarthy, 2002), functionalist cognitive grammar (Nathan, 2008), and others. Indeed, phonological features have withstood almost a century of empirical and theoretical scrutiny and their impact on modern linguistics can hardly be overstated. Their postulation is considered "to be a scientific achievement on the order of the discovery and verification of the periodic table in chemistry" (Jackendoff, 1994), and "the view that speech sounds are not indivisible units but rather are composed of smaller abstract categories called 'distinctive features' is considered to be one of the most important advancements in phonological theory, if not in all of linguistics, of the 20th century" (Mielke & Hume, 2006: 730).

However, in the linguistics and phonetics literature, there is still no agreement on the ontology of features: What are they? What is their nature? Where and in what form do they reside? This lack of consensus can be illustrated by contrasting two pervasive yet mutually incompatible perspectives on the nature of features:

> "Any empirical theory has to have a number of primitives which are definable in terms of concepts which belong outside the theory. In the case of a phonological theory, these are the prime features which are definable in terms of the acoustic or physiological properties of sounds. Each of these features consists of a single measurable property of a kind such that sounds can be said to have this property to a greater or lesser degree. [...] They include features such as Nasality, which is a single measurable property which sounds can have to a greater or lesser degree." (Vennemann & Ladefoged, 1971: 13)

> "Features correspond to controls in the central nervous system which are connected in specific ways to the human motor and auditory systems. In speech perception detectors sensitive to the properties […] are activated, and appropriate information is provided to centers corresponding to the distinctive feature[s] […]. This information is forwarded to higher centers in the nervous system where identification of the utterance takes place. In producing speech, instructions are sent from higher centers in the nervous system to the different feature[s] […] about the utterance to be produced. The features then activate muscles that produce the states and configurations of different articulators." (Halle, 1983/2002: 109–110)

The first definition localizes features at the level of acoustics and articulation, explicitly stating that they are aspects of phonetic substance. The second definition treats features as purely mental units and situates them in the central nervous system, clearly differentiating them from phonetic substance. So, where can a scientist hope to find features? In the movements of the articulators, in the sound waves, in the brain, somewhere else? Clearly, by not determining the ontological status of features, progress is impeded in accounting for the representation and use of spoken language, be it from a phonological, phonetic, psycholinguistic or neurolinguistic perspective (Poeppel & Embick, 2005; Poeppel et al., 2020). Therefore, in order to make these domains more commensurable, this paper addresses two fundamental research questions regarding phonological features: What *kind* of thing is a feature? What is the relationship between features and phonetic substance?

FEATURES AS SYMBOLS IN THE BRAIN

There is a long tradition in phonology that assumes features are abstract mental units that have a lawful but highly indirect relation to phonetic substance:

> "Considerations of this nature [that languages do not make free use of acoustic values or articulatory properties] were much in our minds [...] when Jakobson, Fant and I were working on *Preliminaries to Speech Analysis*, and it was these considerations that led us to draw a sharp distinction between distinctive features, which were abstract phonological entities, and their concrete articulatory and acoustic implementation. Thus, in *Preliminaries* we spoke not of 'articulatory features' or of 'acoustic features', but of 'articulatory and/or acoustic correlates' of particular distinctive features." (Halle, 1983/2002: 94)

As phonology in the 20th century progressed from the taxonomic and mostly anti-mentalist structuralism to the cognitively and neurobiologically oriented generative perspective (Fischer-Jørgensen, 1975: §9; Anderson, 1985: §12; Volenec, 2020: §2), attempts were made to connect feature theory to human neural structures:

> "In articulatory terms each feature might be viewed as information the brain sends to the vocal apparatus to perform whatever operations are involved in the production of the sound, while acoustically a feature may be viewed as the information the brain looks for in the sound wave to identify a particular segment as an instance of a particular sound." (Kenstowicz & Kisseberth, 1979: 239)

In this paper, I propose that continuing this mentalistic tradition and sharpening its claims about the neurobiological reality of phonological features has positive consequences for the study of phonology, phonetics, and their interface. Specifically, I aim to show how features can be productively explored from the perspective of cognitive neuroscience (Poeppel & Embick, 2005; Mildner, 2008; Gallistel & King, 2009; Poeppel, 2012; Embick & Poeppel, 2014; Kemmerer, 2023).

According to Gallistel and King (2009: 72), the atomic elements of mental representations are called symbols. These are “physical entities in a physically realized representational system”, where the physical system in the case of phonological symbols, and all other cognitive symbols, is the human brain. Thus, the working hypothesis here is that phonological features are symbols realized in the human brain. The function of these symbols is to encode, or represent, phonologically relevant aspects of speech. In speech production, they instruct the motor system to carry out the appropriate articulatory movements (Volenec & Reiss, 2017). In speech perception, they are used to decode and parse the incoming utterance into linguistic units (Monahan et al., 2022). The common properties of all neural symbols are (at least) distinguishability, combinability and encoding efficacy.

The standard assumption in cognitive neuroscience is that different symbols are distinguished by place coding of neural activity, rate coding, time coding, or, most likely, some combination of those (Kemmerer, 2023: 8–9). Of course, we are still far from being able to state precisely how features *qua* neural symbols are realized in the brain, but experimental studies are consistently emphasizing the importance of neural circuitry in the superior-most part of the superior temporal gyrus (STG), superior temporal sulcus (STS), and Brodmann (BA) areas 44 and 6 (Figure 1).

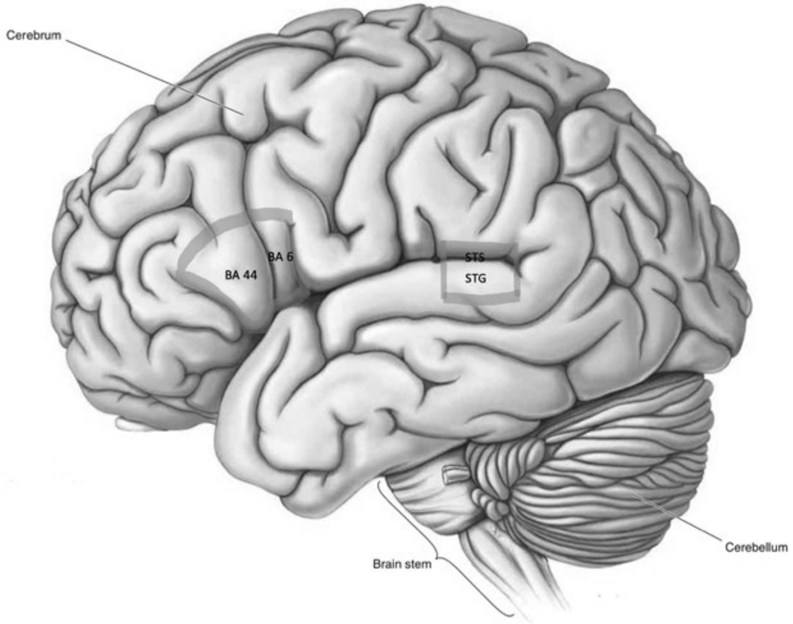

Figure 1. Regions of the human brain (left hemisphere) where phonological features are assumed to be encoded. Brodmann area (BA) 44 is part of Broca’s area. BA 6 is part of the premotor cortex. STS = superior temporal sulcus; STG = superior temporal gyrus. Image adapted from Bear et al. (2016: 222).

The representations of articulatory correlates of features are encoded in the posterior inferior frontal gyrus of the left hemisphere, traditionally known as Broca's area (Okada et al., 2018). More specifically, Hickok (2012: 138) reports that *pars opercularis* (BA 44) and the ventral-most part of BA 6 store articulatory programs needed to reach the auditory targets imposed by features, as depicted in Figure 1. These auditory targets, i.e., the representations of auditory correlates of features, are encoded in the STG and the STS, approximately corresponding to the encircled area on the right in Figure 1. Mesgarani et al. (2014) showed that acoustic phonetic information is represented in the STS and is distributed along five distinct areas, each corresponding to a general 'manner of articulation' parameter. By measuring the responses in implanted electrical cortical grids placed along the superior-most part of the temporal gyrus, they found that one electrode responded *selectively* to stops, one to sibilant fricatives, one to low back vowels, one to high front vowels and a palatal glide, and one to nasals. Similarly, Bouchard et al. (2013) constructed an auditory-based 'place of articulation' cortical map in the STG, confirming labial, coronal and dorsal place features with different electrodes, and cutting across various manner classifications. Using magnetoencephalography (MEG), Scharinger et al. (2012) localized three vowel features – height, frontness and rounding – in different parts of the STG. Using functional magnetic resonance imaging (fMRI) and multivoxel pattern analysis, Arsenault and Buchsbaum (2015) showed that neural representations corresponding to the categorical features of voicing, manner of articulation, and place of articulation are distributed bilaterally throughout the secondary and association areas of the STG.

Features *qua* neural symbols also meet the criterion of combinability. A cornerstone of modern phonology is the notion that features can be grouped into sets to construct higher-level, non-atomic data structures (Chomsky & Halle, 1968; Kenstowicz, 1994; Volenec & Reiss, 2020). An unordered, unstructured set of features constitutes a phonological segment, while a particular organization of segments constitutes a data structure of the next higher level, namely a syllable. This combinability of features allows phonology to construct complex symbols from an inventory of simple parts, and provides an explanation for natural class behavior: different structures can behave alike because they contain identical substructures.

Features are also an efficacious way of encoding information, since their combining leads to combinatoric explosion (Reiss & Volenec, 2022). For example, if we assume that the brain stores and uses only 30 binary features with the possibility of underspecification, then from this small set of primitive symbols we can construct 3^{30} or about 206 trillion different segments. Of course, the richness that arises from feature combinability should not be taken to imply that any particular language should come close to exploiting the full range of possibilities. Instead, what we expect to find in particular languages is in line with the traditional view of feature combination: “No language has as many phonemes as there are possible combinations of the utilized distinctive features” (Halle, 1954: 209). A positive consequence of this combinatoric explosion is that such richness goes a long way toward eliminating the need for a phonetic module of grammar that is specific to each language (as in Keating, 1984, 1990; Scheer, 2020), which simplifies the sequence of conceptual steps needed to account for the externalization of language (see Volenec & Reiss, 2020, for elaboration).

Recent neuropsychological studies have shed light on some other aspects of features that are significant for phonological theory, namely their discreteness, binarity, and potential underspecification. By eliciting magnetic mismatch fields in an oddball paradigm, Phillips et al. (2000) have shown that the left hemisphere STG has access to representations of discrete and binary phonological categories. In other words, their study has “demonstrate[d] the all-or-nothing property of phonological category membership” (Phillips et al., 2000, 1038), where this category membership is determined on the basis of phonological features and not on the basis of general categorical auditory perception. This finding contradicts the claim that gradient articulatory gestures serve as basic units in phonological computation (as in Browman & Goldstein, 1989). Furthermore, Scharinger et al. (2016) found that a less specified vowel compared to its more specified counterpart resulted in stronger activation in the left STS, thus providing some insight into the neural underpinnings of phonological underspecification. Accordingly, Lahiri & Reetz (2010) claim that the phonological forms of words can also be stored in the mental lexicon with featural underspecification, citing evidence from language change, psycholinguistic, and neurolinguistic data.

Even though, as we can see, some progress has been made in discovering the neural reality of features, we are still far from being able to refer to particular features by stating their exact neurobiological substrate, and therefore have to resort to using symbols (labels, names) to refer to symbols. So when we write, for example, “labial”, we use a sequence of letters to form a symbol for a particular feature, which in turn is also a symbol, just in the brain. In other words, “labial” is a non-neural symbol for a neural symbol. We, the researchers, need these phonetic labels to know what we are talking about; the brain does not. The brain does not need such phonetic labels because the transduction algorithms at the phonology–phonetics interface (see section 4 below) interpret the identity of a feature by the place of the neural activity in the brain and its temporal properties (Khalighinejad et al., 2017). This is similar to how a computer does not retrieve the identity of a symbol solely on the basis of its form (1s and 0s), but rather by combining the information about the form with the location and context in the memory (Gallistel & King, 2009: 73). Possibly, the actual form of all features is the same – a neural spike (i.e., an action potential). But more importantly, the unique location of the spike, and the rate of its repetition, is how the transducer determines the identity of the feature and 'knows' which neuromuscular program (e.g., labiality and not, say, nasality) to assign to it. It can of course be debated whether it is misleading or not to use phonetic labels such as "labial" to refer to features *qua* neural symbols, and whether there is a better solution to this (cf. Sheer, 2020: 213). But a decision on this issue has no bearing on the actual nature of features: the neural symbol is, of course, the same irrespective of whether we refer to it as "labial" or by using a non-phonetic label such as "feature 6".

PHONOLOGICAL FEATURES ARE DEVOID OF PHONETIC SUBSTANCE

The symbolic nature of a symbol is that it stands for something *else*, something that is not the same as the symbol. That for which a symbol stands, that which it represents, is variably called its referent, or correlate, or the represented. Phonological features are symbols that refer to aspects of speech. For example, the feature NASAL is a symbol in the brain (a particular neural circuit) that refers to the flowing of air through the nose during speech. Here, it is of utmost importance not to "make the common mistake of confusing the symbol with what it represents" (Gallistel & King, 2009: 56) because "the tendency to confuse symbols with the things they refer to is so pervasive that it must be continually cautioned against" (Gallistel & King 2009: 62). There is a connection between phonological features and speech, but this connection is complex and indirect (see section 4 below), and features do not encode speech-related information in any straightforward way. In linguistics, information related to speech is called phonetic substance (Hjelmslev, 1943; Hale & Reiss, 2000, 2008). It is the totality of the articulatory, acoustic, and auditory properties and processes that constitute speech. For example, properties and processes of speech such as movements of the tongue, values of formants, loudness, duration expressed in milliseconds, etc., fall under the rubric of substance. Since features are symbols physically realized in the brain, they cannot contain phonetic substance. In other words, features are substance-free. Believing that features 'are' substance or that they 'contain' substance (e.g., Ohala, 1990: 162; Flemming, 2001: 9–10) is just an instance of the aforementioned mistake of confusing the symbol with what it represents.

It should also be emphasized that the substance-free nature of features is confirmed by ample evidence from neuroscience. Phillips et al. (2000: 1040) have concluded that “there is good reason to distinguish the acoustic and phonetic representations that underlie categorical perception from the discrete phonological category representations involved in lexical storage and phonological computation,” and that when it comes to phonological computation, “all within-category contrasts are lost: e.g., all different tokens of /d/ are treated by phonological processes as exactly the same”, irrespective of the phonetic substance that is indirectly associated with the bundle of features that we conventionally label as /d/. Magrassi et al. (2015) have shown that the activity of language areas in the brain is organized in terms of features even when language is generated mentally before any utterance is produced or heard, that is, when there is no phonetic substance whatsoever. Similarly, Okada et al. (2018) have conducted an fMRI investigation of silent word sequence production (i.e., the subjects read words in their minds) where the stimuli (different words displayed one after another on a screen) varied in the degree of featural similarity in consonant onset position. The experiment confirmed a featural organization of investigated word sequences in absence of overt speech. These studies suggest that phonological features cannot be equated with the phonetic correlates that are typically associated with them, which is to say that features and phonetic substance are two different things.

As substance-free symbols, features do not contain information on the temporal coordination of muscle contractions, on the spectral configuration of the acoustic target to be reached, and so on. Yet without this information, the respiratory, phonatory, and articulatory systems cannot produce speech. The sensorimotor (SM) system which is in charge of speech production requires information about substance and time in order to arrange the articulatory score (Levelt et al., 1999; Guenther, 2016; Hickok, 2019), so this information has to be integrated into a representation before being fed to the SM system. Therefore, it is necessary to posit a transduction component that connects phonological competence with the vastly different SM system. The theory of that component and the component itself are called Cognitive Phonetics (Volenec & Reiss, 2017; Reiss & Volenec, 2022).

COGNITIVE PHONETICS: FROM SUBSTANCE-FREE FEATURES TO PHONETIC SUBSTANCE

Cognitive Phonetics (CP) is a neurobiologically grounded theory of the phonology–phonetics interface. It proposes that the phonology–phonetics interface consists of at least two transduction procedures that convert the substance-free output of phonological grammar into a representational format that contains substantive information required by the SM system to externalize language through speech.

The inputs to CP are the outputs of phonology, i.e., surface phonological representations (SRs). SRs are strings of segments, each of which is a set of features. Each feature of SRs is transduced and subsequently receives interpretation by the SM system (see Lenneberg et al., 1967: §3 for an historical antecedent). This transduction is carried out by two ‘algorithms’ in the sense of Marr (1982: 23–24). The *paradigmatic transduction algorithm* (PTA) relates a feature (a symbol in the brain) to a motor program which specifies the muscles that need to be contracted in order to produce an appropriate acoustic effect. The *syntagmatic transduction algorithm* (STA) determines the temporal organization of the neuromuscular activity specified by the PTA. In simpler terms, PTA assigns muscle activity to each feature, STA distributes that activity temporally. These transduction algorithms yield an output representation of CP, which then feeds the SM system. The output of CP is called the *phonetic representation* (PR), and it can be defined as a complex array of temporally coordinated motor programs that activate muscles involved in speech production.

The standard generative schema of phonological competence can now be expanded to accommodate the transduction performed by CP, transforming it into a more complete ‘speech chain’ shown in Figure 2.

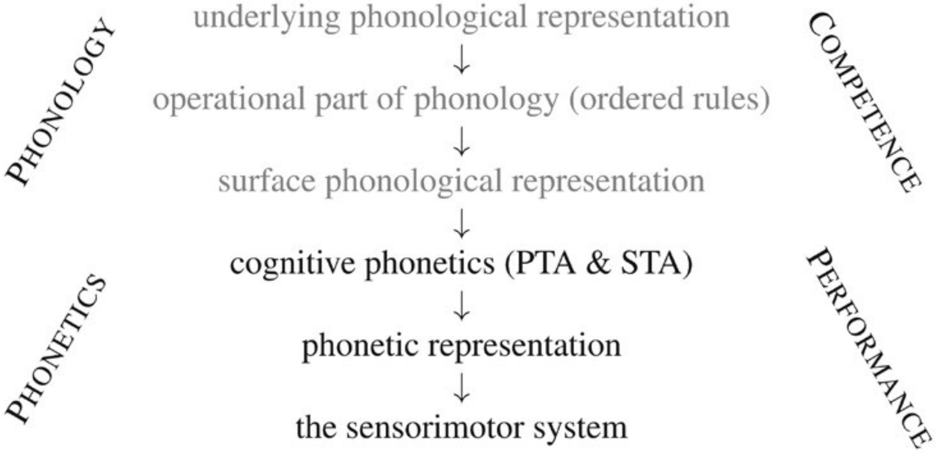

Figure 2. Cognitive Phonetics is a transduction system that connects phonological competence (in gray) and spoken linguistic performance (in black).

The gray parts of the schema represent phonological competence, while the black parts correspond to the initial phonetic steps in speech production. That is, the difference in shading parallels the competence–performance dichotomy: phonology is competence, cognitive phonetics is (one component of) performance. Figure 2 shows that what is loosely referred to in the literature as "the phonetic form" (Chomsky, 1986: 68) or "the externalization of language" (Chomsky et al., 2017: 16) has a sophisticated internal structure.

To clarify the effects of PTA and STA, we can explore in some detail the transduction of a few simple SRs (see Volenec & Reiss, 2019 for further examples). We will see that PTA and STA have considerable implications: they open the possibility of elegantly accounting for subtle yet systematic interactions of two kinds of coarticulatory effects, which is only possible if we assume that the basic units of speech production are indeed transduced phonological features.

Suppose that a language contains SRs [lok] and [luk]. Each segment is a set of features, and vowels [o] and [u] both contain the valued feature +ROUND, on which we will focus. One thing that should be noticed is that [o] and [u] are different in terms of height: [o] is −HIGH, [u] is +HIGH. The PTA takes a segment, scans its feature composition and determines the required muscular activity for the realization of every feature. Roughly, for +ROUND the PTA activates at least four muscles – *orbicularis oris*, *buccinator*, *mentalis*, and *levator labii superioris* (Seikel et al., 2019) – which leads to lip rounding. The difference in PTA's effect on −HIGH and +HIGH is that for the latter, the algorithm raises the tongue body and the jaw, while it does not for the former.

While transducing +ROUND, the PTA takes into account the specification for HIGH and assigns a slightly different lip rounding configuration for [o] than for [u]. Let us refer to a *transduced feature*, which we take to be the basic unit of speech production, as 'PR_{F}', where 'PR' stands for 'phonetic representation' and 'F' stands for an individual valued feature. So, $\mathrm{PR}_{+\mathrm{ROUND}}$ is the transduced feature +ROUND. We can now say that $\mathrm{PR}_{+\mathrm{ROUND}}$ will be different for [o] because of its interaction with $\mathrm{PR}_{-\mathrm{HIGH}}$ than for [u] because of its interaction with $\mathrm{PR}_{+\mathrm{HIGH}}$. Since these interactions involve transduced features within a single segment, [o] or [u], we can refer to these effects as ***intra**segmental coarticulation*. The PTA accounts for intrasegmental coarticulation by assigning a different neuromuscular schema depending on the specification of features from the same segment.

Let us suppose further that, while determining the durational properties of transduced features, the STA temporally extends $\mathrm{PR}_{+\mathrm{ROUND}}$ from the vowel onto the preceding consonant, i.e., in the anticipatory direction. This amounts to the more familiar ***inter**segmental coarticulation* (Volenec, 2015; Mildner, 2018; Liker, 2024), where transduced features from different segments interact. Returning to SRs [lok] and [luk], two things are now apparent: first, $\mathrm{PR}_{+\mathrm{ROUND}}$ is different for [o] than for [u] due to its intrasegmental coarticulation with $\mathrm{PR}_{\mathrm{HIGH}}$; second, [l]'s inherent $\mathrm{PR}_{-\mathrm{ROUND}}$ is now temporally overlapping with the $\mathrm{PR}_{+\mathrm{ROUND}}$ from the adjacent vowels because of intersegmental coarticulation. It is important to note that the difference in $\mathrm{PR}_{+\mathrm{ROUND}}$ from [o] and $\mathrm{PR}_{+\mathrm{ROUND}}$ from [u] will be reflected on the preceding consonant: [l] in [lok] will be articulated differently with respect to lip rounding than [l] in [luk]. Indeed, these differences can clearly be observed in Figure 3. Thus, [l] simultaneously bears the effect of both intra- and intersegmental coarticulation.

Figure 3. Left picture shows the articulation of lip-rounded [l] in [luk]. Right picture shows the articulation of lip-rounded [l] in [lok]. The lip-rounding difference is due to intrasegmental coarticulation inside of the vowels: the realizations of +HIGH and +ROUND are coarticulated in [u], while the realizations of –HIGH and +ROUND are coarticulated in [o]. These different lip-rounding gestures are then extended from the vowels in the anticipatory direction to influence [l], which is a familiar type of intersegmental coarticulation.

CP allows us to account for such subtle yet systematic phonetic variations in an explicit and straightforward way – they follow automatically from PTA and STA, which are independently motivated by the need for transduction.

CP's transduction is universal (non-language-specific) and deterministic, which means that it assigns the same neuro-muscular program to each feature every time that feature is transduced. This also includes all cases of feature combinations that lead to intra- and intersegmental coarticulation. CP thus makes another empirically testable prediction: in principle, given a full and correct list of features, it should be possible to exhaustively describe all possible intra- and intersegmental coarticulatory effects just by using the two algorithms proposed by CP.

It should be stressed that CP's outputs, phonetic representations, should not be equated with actual articulatory movements or with the acoustic output of the human body. What is actually pronounced is further complicated in the process of language externalization by a great number of factors (Chomsky & Halle, 1968: 3; Boeckx, 2010: 78). Transduction is accompanied by other performance factors that have no bearing on either phonology or transduction, factors like muscle fatigue, degree of enunciation, interruptions due to sneezing, trying to achieve a certain intensity level, and many other situational effects, all of which will have an effect on the final output of the body, and will therefore make (co)articulatory variation seem even greater. For that reason, it is not the case that the articulatory and the concomitant acoustic substance will always be identical for each feature or feature combination. However, this apparent "lack of invariance" (Applebaum, 1996: 1541) in the realization of a cognitively invariant category is not a matter of transduction, but rather is a result of accidental performance factors.

A BRIEF OUTLINE OF THE NEUROBIOLOGICAL FOUNDATION OF COGNITIVE PHONETICS

In this section, I will briefly state the neural circuits that correspond to phonological features, to the transduction operations of CP, and to the realization of transduced features by the primary motor cortex. These neural mechanisms are graphically represented in Figure 4.

The activity in parts of the inferior frontal gyrus (IFG) corresponds to the representations of the articulatory correlates of features (Eickhoff et al., 2009; Hickok, 2012), while the activity in parts of the superior temporal gyrus (STG) and sulcus (STS) corresponds to the representations of the auditory correlates of features (Hickok & Poeppel, 2007: 398; Bouchard et al., 2013). An area in the Sylvian fissure at the boundary between the parietal and the temporal lobe (Spt) unifies these two aspects into a complete symbol, a feature (Hickok et al., 2009; Gow, 2012). The symbols are transmitted to the anterior insula (Dronkers, 1996; Blumstein & Baum, 2016) where the PTA is carried out, and to the cerebellum and the basal ganglia (Jueptner & Krukenberg, 2001; Ackerman et al., 2007) where the STA is carried out. The PTA and the STA are integrated in the anterior part of the supplementary motor area (pre-SMA) (Alario et al., 2006; Bohland et al., 2010) to form the phonetic representation, which is a set of neural signals that the primary motor cortex (PMC) sends to the effectors that produce speech.

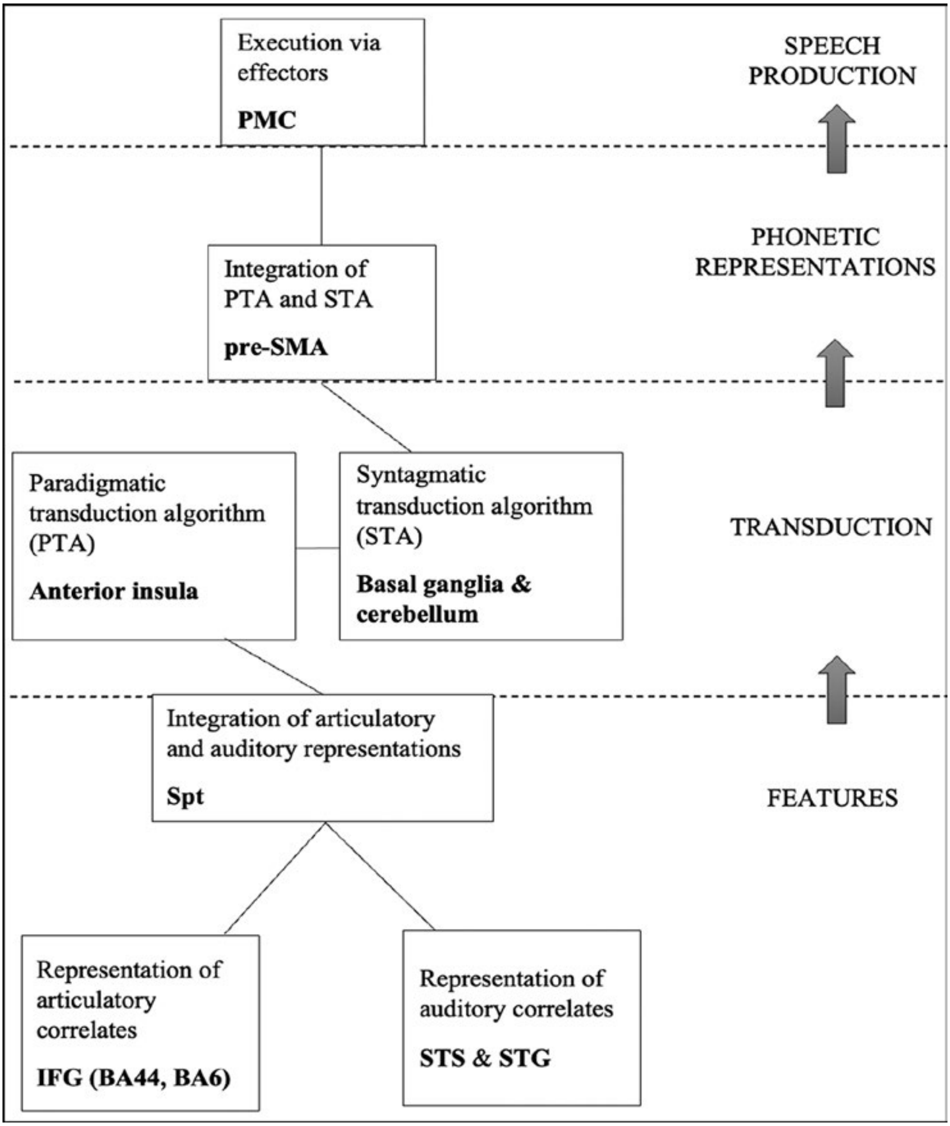

Figure 4. The neurobiological basis of phonological features and of the phonology-phonetics interface, as proposed by the theory of Cognitive Phonetics. IFG = inferior frontal gyrus; BA = Brodmann area; STS = superior temporal sulcus; STG = superior temporal gyrus; Spt = a cortical structure in the posterior part of the Sylvian fissure at the boundary between the parietal and temporal lobes; PTA = paradigmatic transduction algorithm; STA = syntagmatic transduction algorithm; pre-SMA = anterior part of the supplementary motor area; PMC = primary motor cortex.

CONCLUSION

In this paper, adopting a cognitive neuroscience perspective, I proposed that features should be conceptualized as neural symbols with three properties: distinguishability, combinability and encoding efficacy. The function of these symbols is to yield mental representations of phonologically relevant aspects of speech. In speech production, they instruct the motor system to carry out the appropriate articulatory movements, while in speech perception, they are used to decode and parse incoming utterances into linguistic units.

I also outlined a neurolinguistic model of the phonetics–phonology interface called Cognitive Phonetics (CP), which describes how features relate to phonetic substance during speech production. CP proposes that each feature that comprises a surface phonological representation activates speech movements via two simple neural procedures: the paradigmatic transduction algorithm and the syntagmatic transduction algorithm. I showed how CP can account not only for the well-known *inter*segmental coarticulation, but also for the previously less studied *intra*segmental coarticulation, suggesting that the basic units of speech production are transduced features.

CP, as presented in this paper, has several limitations that need to be addressed in future work: How exactly does the kinematic implementation of PTA and STA work? How does CP relate to (and to what extent is it compatible with) other neurobiologically grounded theories of speech production such as DIVA? How can CP account for those patterns that have been deemed as examples of so-called ‘language-specific phonetics’? How do we arrive at a stage where we can actually identify particular neural circuits that correspond to phonological features?

REFERENCES

Ackermann, H., Mathiak, K., & Riecker, A. (2007). The contribution of the cerebellum to speech production and speech perception: clinical and functional imaging data. *The Cerebellum*, 6(3), 202–213.

Alario, X., Chainay, H., Lehericy, S., & Cohen, L. (2006). The role of the supplementary motor area (SMA) in word production. *Brain Research*, 1076(1), 129–143.

Anderson, S. R. (1985). *Phonology in the Twentieth Century. Theories of Rules and Theories of Representations*. Chicago: The University of Chicago Press.

Appelbaum, I. (1996). The lack of invariance problem and the goal of speech perception. In H. T. Bunnell & W. Idsardi (Eds.), *Proceedings of the 4th*

International Conference on Spoken Language Processing (pp. 1541–1544). Philadelphia: IEEE.

Arsenault, J. S., & Buchsbaum, B. R. (2015). Distributed neural representations of phonological features during speech perception. *Journal of Neuroscience*, 35(2), 634–642.

Bear, M. F., Connors, B. W. & Paradiso, M. A. (2016). *Neuroscience – Exploring the Brain*. Philadelphia: Wolters Kluwer.

Blumstein, S. E., & Baum, S. R. (2016). Neurobiology of speech production: Perspective from neuropsychology and neurolinguistics. In G. Hickok & S. L. Small (Eds.), *Neurobiology of Language* (pp. 689–699). London: Elsevier.

Boeckx, C. (2010). *Language in Cognition. Uncovering Mental Structures and the Rules Behind Them.* Malden: Wiley-Blackwell.

Bohland, J. W., Bullock, D., & Guenther, F. H. (2010). Neural representations and mechanisms for the performance of simple speech sequences. *Journal of Cognitive Neuroscience*, 22(7), 1504–1529.

Bouchard, K. E., Mesgarani, N., Johnson, K., & Chang, K. (2013). Functional organization of human sensorimotor cortex for speech articulation. *Nature*, 495(7441), 327–332.

Browman, C. P., & Goldstein, L. (1989). Articulatory gestures as phonological units. *Phonology*, 6(2), 201–251.

Chomsky, N. (1986). *Knowledge of Language. Its Nature, Origins, and Use*. New York: Praeger.

Chomsky, N., & Halle, M. (1968). *The Sound Pattern of English*. New York: Harper & Row.

Chomsky, N., Gallego, Á. J., & Ott, D. (2019). Generative grammar and the faculty of language: Insights, questions, and challenges. *Catalan Journal of Linguistics*, 229–261.

Dronkers, N. F. (1996). A new brain region for coordinating speech articulation. *Nature*, 384(6605), 159–161.

Eickhoff, S. B., Heim, S., Zilles, K., & Amunts, K. (2009). A systems perspective on the effective connectivity of overt speech production. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 367(1896), 2399–2421.

Embick, D., & Poeppel, D. (2014). Towards a computational(ist) neurobiology of language: Correlational, integrated and explanatory neurolinguistics. *Language, Cognition and Neuroscience*, 30(4), 357–366.

Fischer-Jørgensen, E. (1975). *Trends in Phonological Theory. A Historical Introduction*. Copenhagen: Akademisk Forlag.

Flemming, E. (2001). Scalar and categorical phenomena in a unified model of phonetics and phonology. *Phonology*, 18(1), 7–44.

Gallistel, C. R., & King, A. P. (2009). *Memory and the Computational Brain: Why Cognitive Science Will Transform Neuroscience*. Malden: Wiley-Blackwell.

Gow, D. W. (2012). The cortical organization of lexical knowledge: a dual lexicon model of spoken language processing. *Brain and Language*, 121(3), 273–288.

Guenther, F. H. (2016). *Neural Control of Speech*. Cambridge: MIT Press.

Halle, M. (1954). The strategy of phonemics. *Word*, 10(2–3), 197–209.

Halle, M. (1983/2002). On distinctive features and their articulatory implementation. In M. Halle, *From Memory to Speech and Back* (pp. 105–121). Berlin – New York: De Gruyter Mouton.

Hale, M., & Reiss, C. (2000). 'Substance Abuse' and 'Dysfunctionalism': Current Trends in Phonology. *Linguistic Inquiry*, 31(1), 157–169.

Hale, M., & Reiss, C. (2008). *The Phonological Enterprise*. Oxford: Oxford University Press.

Hickok, G. (2012). Computational neuroanatomy of speech production. *Nature Reviews Neuroscience*, 13(2), 135–145.

Hickok, G. (2019). The Dorsal Stream Auditory-Motor Interface for Speech. In G. Zubicaray & N. Schiller (Eds.), *The Oxford Handbook of Neurolinguistics* (pp. 498–515). Oxford: Oxford University Press.

Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews Neuroscience*, 8(5), 393–402.

Hjelmslev, L. (1943/1969). *Prolegomena to a Theory of Language*. Madison: University of Wisconsin Press.

Jackendoff, R. (1994). *Patterns in the Mind. Language and Human Nature*. New York: BasicBooks.

Jakobson, R. (1939). Observations sur le classement phonologique des consonnes. J. R. Joos (Ed.), *Proceedings of the 3rd International Congress of Phonetic Sciences* (pp. 34–41). Ghent: Imprimerie Charles Bulens.

Jakobson, R. (1962). *Selected Writings: Phonological Studies*. The Hague: Mouton.

Jakobson, R., & Halle, M. (1956). *Fundamentals of Language*. The Hague: Mouton.

Jakobson, R., Fant, G., & Halle, M. (1952). *Preliminaries to Speech Analysis.* Cambridge: MIT Press.

Jueptner, M., & Krukenberg, M. (2001). Motor system: Cortex, basal ganglia, and cerebellum. *Neuroimaging Clinics of North* America, 11(2), 203–219.

Keating, P. (1984). Universal phonetics and the organization of grammars. *UCLA Working Papers in Phonetics*, 59, 35–49.

Keating, P. (1990). Phonetic representations in a generative grammar. *Journal of phonetics*, 18(3), 321–334.

Kemmerer, D. (2022). *Cognitive Neuroscience of Language*. New York – London: Routledge.

Kenstowicz, M. (1994). *Phonology in Generative Grammar*. Oxford: Blackwell.

Kenstowicz, M., & Kisseberth, C. (1979). *Generative Phonology. Description and Theory.* New York: Academic Press.

Khalighinejad, B., da Silva, G. C., & Mesgarani, N. (2017). Dynamic encoding of acoustic features in neural responses to continuous speech. *Journal of Neuroscience*, 37(8), 2176–2185.

Lahiri, A., & Reetz, H. (2010). Distinctive features: Phonological underspecification in representation and processing. *Journal of Phonetics*, 38(1), 44–59.

Lenneberg, E. (1967). *Biological Foundations of Language*. New York: Wiley.

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, 22(1), 1–38.

Liker, M. (2024). *Koartikulacija – što sve ne znamo o govoru?* Zagreb: Ibis grafika.

Magrassi, L., Aromataris, G., Cabrini, A., Annovazzi-Lodi, V., & Moro, A. (2015). Sound representation in higher language areas during language generation. *Proceedings of the National Academy of Sciences*, 112(6), 1868–1873.

Marr, D. (1982/2010). *Vision. A Computational Investigation into the Human Representation and Processing of Visual Information.* Cambridge: MIT Press.

McCarthy, J. J. (2002). *A Thematic Guide to Optimality Theory*. Cambridge: Cambridge University Press.

Mesgarani, N., Connie, C., Johnson, K., & Chang, E. F. (2014). Phonetic feature encoding in human superior temporal gyrus. *Science*, 343(6174), 1006–1010.

Mielke, J., & Hume, E. (2006). Distinctive Features. In K. Brown (Ed.), *Encyclopedia of Language and Linguistics*, Vol. 2, 723–731. Philadelphia: Elsevier.

Mildner, V. (2008). *The Cognitive Neuroscience of Human Communication*. New York – London: Lawrence Erlbaum Associates.

Mildner, V. (2018). Aspects of coarticulation. In M. Gosy (Ed.), *Challenges in Analysis and Processing of Spontaneous Speech* (pp. 27–48). Budapest: MTA.

Monahan, P. J., Schertz, J., Fu, Z., & Pérez, A. (2022). Unified coding of spectral and temporal phonetic cues: Electrophysiological evidence for abstract phonological features. *Journal of cognitive neuroscience*, 34(4), 618–638.

Nathan, G. S. (2008). *Phonology: A Cognitive Grammar Introduction*. Amsterdam: John Benjamins Publishing.

Ohala, J. J. (1990). There is no interface between phonology and phonetics: a personal view. *Journal of Phonetics*, 18(2), 153–171.

Okada, K., Matchin, W., & Hickok, G. (2018). Neural evidence for predictive coding in auditory cortex during speech production. *Psychonomic Bulletin and Review*, 25(1), 423–430.

Phillips, C., Pellathy, T., Marantz, A., Yellin, E., Wexler, K., Poeppel, D., McGinnis, M., & Roberts, T. (2000). Auditory cortex accesses phonological categories: An MEG mismatch study. *Journal of Cognitive Neuroscience*, 12(6), 1038–1055.

Poeppel, D. (2012). The maps problem and the mapping problem: Two challenges for a cognitive neuroscience of speech and language. *Cognitive Neuropsychology*, 29(1–2), 34–55.

Poeppel, D., & Embick, D. (2005). Defining the relation between linguistics and neuroscience. In A. Cutler (Ed.), *Twenty-First Century Psycholinguistics: Four Cornerstones* (pp. 103–118). London: Psychology Press.

Poeppel, D., Mangun, G. R., & Gazzaniga, M. (Eds.). (2020). *The Cognitive Neurosciences*. Cambridge: MIT Press.

Prince, A., & Smolensky, P. (1993/2004). *Optimality Theory: Constraint Interaction in Generative Grammar*. Oxford: Blackwell.

Reiss, C., & Volenec, V. (2022). Conquer primal fear: Phonological features are innate and substance-free. *Canadian Journal of Linguistics*, 67(4), 581–610.

Scharinger, M., Monahan, P. J., & Idsardi, W. J. (2012). Asymmetries in the processing of vowel height. *Journal of Speech, Language, and Hearing Research*, 55(3), 903–918.

Scharinger, M., Domahs, U., Klein, E., & Domahs, F. (2016). Mental representations of vowel features asymmetrically modulate activity in superior temporal sulcus. *Brain and Language*, 163, 42–49.

Scheer, T. (2020). On the lexical character of intermodular communication. *Radical: A Journal of Phonology*, 1, 183–239.

Seikel, J., Drumright, D. G., & Hudock, D. J. (2019). *Anatomy and Physiology for Speech, Language, and Hearing.* San Diego: Plural Publishing.

Vennemann, T., & Ladefoged, P. (1971). Phonetic features and phonological features. *UCLA working papers in phonetics*, 21(21), 13–24.

Volenec, V. (2015). Coarticulation. In J. Davis (Ed.), *Phonetics: Fundamentals, Potential Applications and Role in Communicative Disorders* (pp. 47–86). New York: Nova Science Publishers.

Volenec, V. (2020). *Generativna fonologija*. Zagreb: Ibis grafika.

Volenec, V., & Reiss, C. (2017). Cognitive Phonetics: The transduction of distinctive features at the phonology–phonetics interface. *Biolinguistics*, 11, 251–294.

Volenec, V., & Reiss, C. (2019). The intervocalic palatal glide in Cognitive Phonetics. M. Baird & J. Pesetsky (Eds.), *Proceedings of the 49th Meeting of the North East Linguistics Society* (pp. 255–264). Ithaca: Cornell University Press.

Volenec, V., & Reiss, C. (2020). Formal generative phonology. *Radical: A Journal of Phonology*, 2, 1–148.