

**2020**

**Volenec, Veno  
Reiss, Charles**

# **Formal Generative Phonology**

---

Volenec, Veno & Reiss, Charles. 2020. "Formal Generative  
Phonology". Manuscript available in *Radical: A Journal of  
Phonology*.

# FORMAL GENERATIVE PHONOLOGY

VENO, VOLENEC & CHARLES REISS (CONCORDIA UNIVERSITY)

This paper outlines a program for the study of phonology as a branch of cognitive science. Building on the legacy of classical generative phonology and biolinguistics, it provides a theoretical framework that strictly differentiates phonological competence from aspects of articulation, acoustics and perception. We argue that phonological competence is to be characterized as a formal—that is, explicit, logically precise, and substance-free—manipulation of abstract symbols. We propose that a productive way to execute this program is to adopt a model called *Logical Phonology*, where phonological competence is described and explained by a system that maps between phonological data structures like strings via rules constructed from basic set-theoretic operations. We show the merits of this model by applying it to Turkish, Hungarian and English data. The remote and complex relationship between phonological competence and speech is elucidated by *Cognitive Phonetics*, which proposes that the outputs of phonology are transduced via two algorithms into temporally distributed neuromuscular activities. Taken together, *Logical Phonology* and *Cognitive Phonetics* aim to explain the nature of what is loosely referred to in the literature as ‘the externalization of language’ and to delineate its components.

*generative phonology, I-language, substance-free phonology, rule, feature*

## INTRODUCTION

The externalization of language—the conversion of the output of a syntactic derivation to a spoken utterance—is a process of great complexity. Because of the assumed autonomy of syntax from the components of this conversion (Chomsky et al. 2017: 18), the syntactic literature understandably lumps the whole process under a variety of interchangeable and somewhat vague labels such as ‘AP’ for the ‘articulatory-perceptual system’; ‘EXT’ for ‘externalization’; ‘SM’ for the ‘sensorimotor system’; ‘PF’ for either ‘phonological form’ or ‘phonetic form’; or a catch-all ‘PHON’.<sup>1</sup> Syntacticians assume that it is the job of the ‘sound’ researchers, phonologists and phoneticians, to explain what is going on between the output of syntax and the mouth. ‘Sound’ researchers usually do agree that both phonetics and phonology are needed to account for the externalization of syntax, but there is very little agreement about how to characterize these domains and where to place the boundary between them. Furthermore, the theoretical assumptions of various scholarly communities—syntacticians, phonologists and phoneticians—are so mutually incompatible as to preclude a coherent account of ‘EXT’. For example, nativist assumptions based on the Argument from the Poverty of the Stimulus and other considerations widely accepted in the syntactic literature have been increasingly ignored or rejected by phonologists (e.g., Archangeli & Pulleyblank 2015), despite parallel arguments in the two domains. We also find a chasm between phonologists, who tend to accept the necessity of feature-based analyses (whether or not they are nativists with respect to features), and scholars in phonetics and speech processing, who largely reject or ignore phonological features. This situation is comparable to Pylyshyn’s description of the

<sup>1</sup> This claim should not be taken to imply that there cannot be any interaction between phonology and morphosyntax, but rather that from the syntactic output to the acoustic output of the body, there are complexities that are sometimes obscured in the syntactic literature, for example by calling everything from the output of syntax to phonetics ‘the SM system’ or ‘the phonetic form’. To constrain the exposition in this paper, we will abstract away from possible interactions between these grammatical systems, interactions which are covered, for example, in *The Sound Pattern of English* (Chomsky & Halle 1968), more recently by Distributed Morphology (Halle & Marantz 1993; Harley & Noyer 1999), and by other work.

complexities that arise when different research communities try to arrive at unifiable models:

It turns out that the cognitive theorist enjoys considerable latitude in deciding which functions to attribute to the transducer and which to the symbolic processor. The boundary between transducer and symbolic processor can be shifted about, but this shifting carries with it profound implications for cognitive theory. It is possible to assign the deepest mysteries of perception and cognition to different places, and consequently to make the part one is concentrating on at the moment appear either surprisingly simple or unfathomably complex. [...] Cognitive processes can also be simplified by attributing complexity to the functioning of the transducer, but a price must be paid for this shift. The price [...] is that one can trivially explain anything that way. Unless what counts as transduction is constrained in a principled manner, the simplification of the problems of perception gained by postulating certain transducers has, as Bertrand Russell is reported to have said once, all the virtues of theft over honest toil. (Pylyshyn 1984: 148)

There is a high degree of fragmentation currently in phonology and the rapid proliferation of new theories is matched only by their mutual incommensurability. Even theories labeled as ‘generative’ tend to be completely incompatible. As we will argue below, we think that the main factor that contributes to such a fragmentation is a serious confusion about the object of study in phonology, i.e., a misconception about what it is that phonological theory should explain.<sup>2</sup> Of course, such a state of affairs makes it much more difficult to say something coherent about the relationship between phonology and

---

<sup>2</sup> In the introduction to *The Routledge Handbook of Phonological Theory* (2018), Hannahs and Bosch offer some more specific reasons for this fragmentation: “Perhaps the most obvious proximate cause of disagreement in generative phonology is the advent of Optimality Theory (OT). [...] The popularity of OT among a large group of phonologists, alongside its failure to persuade or engage others, may be said to be responsible for some of the ensuing fragmentation within the field of phonology even among those phonologists who consider themselves generativists, in the sense of Chomsky (1957, 1965)” (p. 2). Additionally, they cite issues related to abstractness, increased interest in the sociolinguistic aspects of phonology, and the resurgence of empiricist methods as further contributing factors.

domains related to it, particularly phonetics, thus precluding a principled account of ‘EXT’. The present paper aims to mitigate this fragmentation, first by defining the object of phonological study in a way that is consistent with the philosophical foundations of generative linguistics (Chomsky 1968; 1986; 2000), and then by outlining a research program for the study of that object.<sup>3</sup>

In the next section, we explore the theoretical and methodological consequences of adopting the position that the object of phonological study is completely internal to an individual’s mind/brain. In the rest of the paper we then establish that what is loosely called ‘externalization’ in the literature consists of two parts: (1) *phonology* – a grammatical subsystem governed by rules; (2) *universal phonetics* – a non-grammatical (and therefore *non-language-specific*) component that interprets the outputs of phonology and produces speech. In other words, we decompose ‘EXT’ into a ‘symbolic processor’ and a ‘transducer’, to use Pylyshyn’s (1984) terms. Section 2 sketches *Substance Free Phonology*, a formal theoretical framework that strictly differentiates phonological competence from all aspects of articulation, acoustics and perception. In section 3, we outline a working model for substance-free phonology called *Logical Phonology*, which uses basic set-theoretical notions to describe the representational and computational aspects of phonological competence. *Cognitive Phonetics*, our conception of the phonology-phonetics interface, is presented in section 4. Building on the work of Lenneberg, Marr, Pylyshyn, Halle and others, it proposes a universal transduction system, whose inputs, crucially, are the feature-based surface representations of generative phonology. Summary and conclusions are given in section 5.

---

<sup>3</sup> The main rationale for such a broad goal is that arguing about more specific phonological problems and their explanations is highly unproductive unless we at least agree about what constitutes ‘a phonological problem’ (as opposed to, say, a phonetic one). For example, we think it is not particularly useful to argue about phonological gradience unless we first agree on *where and in what form* gradience is supposed to reside. In the mind? In the movement of the articulators? In the acoustic signal? In all of these domains? Somewhere else? Once we have clearly defined our object of study, we will be in a better position to address such specific problems productively (for example, see section 1.1 where we return to the issue of gradience after defining the object of our study).

While the present study shares its theoretical orientation with substance-free phonology of Hale & Reiss (2008) and related work, particularly in the form of looking at phonology from the I-language perspective, it goes beyond that work and offers the following novel ideas:

- we use the cognitive neuroscience framework outlined in Gallistel & King (2010) to look at phonological features in a neurobiologically realistic way (§2);
- we draw a parallel with amodal completion in vision studies (Michotte & Burke 1951) to apply the Argument from the Poverty of the Stimulus to phonology and argue for the innateness of features (§2);
- we provide arguments against recent influential work (e.g., Mielke 2008; Archangeli & Pulleyblank 2015) that makes problematic claims about the nature of features and rules (§2);
- we use set theory to model phonological representations and computations, exploring the use of priority union in derivations, an idea we apply to natural language data from Hungarian and Turkish (§3);
- we apply David Marr's (1982) insights on the nature of information-processing systems to construct a phonology-phonetics interface theory (§4);
- we argue that our 'Logical Phonology + Cognitive Phonetics' combination eliminates the need for a language-specific phonetics module in the grammar, thus simplifying the sequence of conceptual steps needed to account for the externalization of language (§4).<sup>4</sup>

---

<sup>4</sup> This paper is clearly programmatic in nature, and as such, it does not argue *directly* against approaches to phonology, phonetics or their interface that differ from those that are proposed here. The reader should also keep in mind that some of our arguments—for example, for the innateness of features—may be relevant to selecting between, say, competing versions of Optimality Theory or Government Phonology. Our claim is not that the argument we give leads inexorably to our 'Logical Phonology + Cognitive Phonetics' model, but rather that we have made some progress in defining a coherent split between phonetics and phonology and also provided a sketch of a promising model of their interface.

## 1 PHONOLOGY AS A BRANCH OF COGNITIVE SCIENCE

### 1.1 THE OBJECT OF STUDY IN PHONOLOGY: I-LANGUAGE

George Miller (1998) noted that the most profound consequence of the cognitive revolution of the late 1950s was the realization that in any branch of cognitive psychology, including linguistics, observable behavior, such as spoken utterances, is evidence, and not the object of study. The actual objects of study are the cognitive systems that underlie observable behavior. The utterances are merely one kind of evidence for the cognitive systems.<sup>5</sup> To put it in the bluntest terms, then, phonology is *not* the study of speech, but speech provides evidence for the study of phonological competence.

Phonology was included in the cognitive revolution right from the outset, in recognition of the crucial distinction between its object of study (phonological competence) and its main source of evidence (speech production and perception) (Chomsky 1951, 1957a, 1957b, 1964; Chomsky *et al.* 1956; Halle 1959, 1962, 1964). However, despite the clarity of this early work, phonology has not been consistently treated as a branch of cognitive science. Instead, there have always been strands of work that adopt the formal trappings of generative phonological theory without adopting its mentalist commitments. This attitude seems to stem from both a failure to engage with the philosophical foundations of the cognitive revolution, and also from a related preoccupation with describing linguistic *behavior* and not the unobservable mental competence that makes that behavior possible. This preoccupation is undeniably gaining in popularity in recent years,<sup>6</sup> as

<sup>5</sup> Miller's (1998) observation should not be understood as suggesting that cognitive science is different from other sciences in this respect, although there is some confusion about this matter. For example, Peter Norvig, former Director of Research at Google, has specifically critiqued the fact that in generative linguistics the object of study is *not* a set of observations: "I can't imagine Laplace saying that observations of the planets cannot constitute the subject-matter of orbital mechanics, or Maxwell saying that observations of electrical charge cannot constitute the subject-matter of electromagnetism" (quoted from <http://norvig.com/chomsky.html>). We think, on the contrary, that neither Laplace nor Maxwell would describe their own *observations* as the object of study. Observations are subject to margins of error, faulty equipment, observer bias, and so on. The laws of nature are not.

<sup>6</sup> This trend has recently been characterized as "[t]he present regrettable return of neo-empiricism, in the

witnessed by the pervasiveness of output-driven models of phonology and by the prevalent orientation towards ‘the surface’ throughout the field. What muddles this issue even further is the fact that a vast amount of phonological work is either indifferent or agnostic about whether phonology is concerned with the mind/brain or with sound patterns that just exist ‘out there’.

Since the reasons for the paradigm shift in the late 1950s were to our knowledge never disputed, we see no reason *not to* adopt strict mentalism in phonology and accept *all* accompanying philosophical and methodological commitments that come with treating phonology as a branch of cognitive science. In the rest of this section we will therefore provide a mentalistic definition of our object of study and explore the methodological implications that such a definition has for phonology.

Our object of study, *phonological competence*, is an aspect of *I-language* (Chomsky 1986). Here, the prefix *I-* stands for several key properties that characterize this technical notion of language. First, language is *internal* to the human mind. Since the mind is just the functioning of the brain viewed from a sufficiently abstract point of view, language can ultimately be regarded as a biological object—a certain subsystem of the brain. Consistent with the conceptual shift Miller referred to, the object of our study is fully internal to a speaker and not something ‘out in the world’, belonging to a ‘speech community’ in some vague sense. Second, following directly from the first, language is a property of an *individual*. Since no two individuals have exactly the same experiences during the period of language acquisition, it is not surprising that they end up with *I-languages* that differ from each other. To the extent that they are similar, language-based communication between individuals occurs with more or less facility. The question also arises of why, despite these differences of experience, individuals may acquire very similar *I-languages*. In other words, by recognizing *I-language* as the object of study, we can

---

shape of Bayesian learning models, statistical regularities, collective exchanges and the extraction of generic cognitive patterns, recurrent networks, and similar notions.” (Piattelli-Palmarini 2018: 211)



start to ask what factors determine the plasticity of the human language faculty. Third, language is an *intensional* system, which means that it is properly characterized by means of a finite set of precise statements that describe the generative capacity of that system. Language characterizes, in terms of rules or functions, the members of an unbounded set of structured expressions; these members are not listed in long term memory. In every I-language there is an infinity of members (sentences), so defining the system extensionally, by listing all of its members, is impossible.

Additionally, language is a kind of *implicit* or *tacit knowledge*. Its content is predominately below the level of consciousness and beyond the reach of introspection. The main task of linguistics is to provide an explicit characterization of this implicit knowledge system in terms of a framework that can account for all possible I-languages.<sup>7</sup> The speakers' judgments about their own linguistic output can be considered an important source of evidence for discovering the nature of this knowledge, but there is no reason to think that a speaker can report perfectly (i.e., completely reliably) on the contents of their I-language.

Crucially, the term *I-language* does not include the notion *innate*, although *nativism*, the existence of a non-trivial genetically determined component of the human language faculty, has also been a tenet of most proponents of the I-language perspective. In brief, one could imagine a cognitive psychologist who accepts that humans have I-languages encoded in their brains based on some kind of all-purpose pattern discovery capacity, without any genetically determined substrate that is specific to language. For such a psychologist, an I-language would be parallel to a person's knowledge of the rules of chess. So, I-language does *not* imply nativism. However, there *is* an implication in the other direction: if one accepts the existence of an innate aspect of the language faculty, then one also accepts, at a minimum, the internalist part of the I-language perspective,

---

<sup>7</sup> "The ultimate outcome of these investigations should be a theory of linguistic structure in which the descriptive devices utilized in particular grammars are presented and studied abstractly, with no specific reference to particular languages." (Chomsky 1957: 11)

because whatever is innate must also be internal to an individual's brain. It follows, then, that a rejection of internalism entails a rejection of nativism. This logic will be important as we demonstrate, first, that nativism is not universally accepted by phonologists, and second, that the rejection of nativism can be tied to a more or less explicit rejection of a crucial aspect of internalism (see sections 2 and 3).

It is of paramount importance to distinguish—both in theory and in practice—between a cognitive system and the use of that system for a particular purpose on a particular occasion. That is, we must distinguish between I-language and the *use* of I-language. To be able to do that clearly, linguistics provides the well-known dichotomy between *linguistic competence* (an I-language) and *linguistic performance* (the use of I-language in, say, communication). Linguistic competence is always a core factor in any act of linguistic performance, interacting with many other linguistically irrelevant factors such as a speaker's level of fatigue, creativity, having a sore throat, free will, or being distracted by a mobile phone. Competence is present even when it cannot be used in acts of performance, as demonstrated, for example, by numerous studies documenting recovery from post-stroke aphasia (Pinker 1994; Hillis 2010; Tippet & Hillis 2016), and by infant comprehension studies which show that children understand speech at a level that far exceeds their ability to produce it (Hirsh-Pasek & Golinkoff 1996; Lust 2006; Smith 2010; Bergelson & Swingley 2012). Since competence is always at the core of every performance act, but crucially not vice versa, we can say that competence has *epistemological priority* over performance: We must first learn about the nature of competence in order to be able to undertake a scientific study of performance.

Language (competence, I-language) is a module of the human mind.<sup>8</sup> A *module* is

---

<sup>8</sup> Language is modular in two ways: It is functionally separable from other, non-linguistic aspects of the mind, and it itself consists of (sub)modules. The main source of evidence for the modular nature of a cognitive capacity is *double dissociation*: If it can be demonstrated that system A can be impaired without affecting the functioning of system B, and if, on a separate occasion, it can be demonstrated that system B can be impaired without affecting the functioning system A, then systems A and B constitute distinct modules. There is a massive amount of empirical evidence of this kind pertaining to the modularity of language. Language is doubly dissociated from acoustic processing (Zatorre et al. 1992;

an informationally encapsulated cognitive system governed by principles not completely shared by any other module. In line with the computational-representational conception of the mind, each module consists of a finite set of primitive, atomic symbols (basic representations) and a finite set of functions that manipulate those symbols (basic computations). As we will show in more detail in section 2, we can also make concrete hypotheses about the symbols and functions of phonological competence.

Different modules communicate via *interfaces*: configurations in which the outputs of one module serve as the inputs to another module. Manipulation (e.g., reordering, regrouping, deletion, addition) of representational elements within a module, and without a change in the representational format is called *computation*. Computation occurs within a single module. Conversion of an element in one form into a distinct form, i.e., a mapping between dissimilar formats is called *transduction*. Transduction occurs at the interface between different modules, and it is a way in which modules communicate with each other. So, modules compute, interfaces transduce. Computation within the phonological module is described in sections 2 and 3, while transduction of information between the phonological module and the system for speech production is described in section 4 on Cognitive Phonetics.

Since language is a module of the mind and therefore a subsystem of the brain,

---

Burton et al. 2000; Phillips et al. 2000; Dehaene-Lambertz & Pena 2001; Dehaene-Lambertz et al. 2002), from attention (Pulvermüller et al. 2008; MacGregor et al. 2012; Blank et al. 2014), from visual and spatial cognition (Emmorey et al. 1993; Wang et al. 1995; McGuire et al. 1997; Emmorey 2002), from ‘general intelligence’ (Lenneberg 1967; Rondal 1995; Curtiss 1982, 1988a, 1988b, 1995, 2011; Smith & Tsimpli 1995; Smith et al. 2010); from non-linguistic communication (Corina et al. 1992; Willems et al. 2009; Grosvald et al. 2012); from arithmetic competence (Schaller 1995; Grinstead et al. 1997; Brannon 2005; Curtiss 2011); from social cognition (Karmiloff-Smith et al. 1995; Smith 2003; Smith et al. 2010; Gernsbacher et al. 2016), from pragmatics (Bottini et al. 1994; Champagne-Lavau & Joannette 2009). In all of these cases, a double dissociation between language and other cognitive systems has been experimentally demonstrated. Furthermore, there is ample evidence for the separability of systems *within* linguistic competence: for the lexicon (Hart et al. 1985; Caramazza 1988; Jodzio et al. 2008); for semantics (Binder et al. 2009; Skeide et al. 2014; Patterson & Ralph 2016); for phonology (Phillips et al. 2000; Leonard 2017); and for syntax (Grodzinsky 1986; Grodzinsky & Finkel 1998; Friedmann et al. 2006; Bastiaanse & Zonnefeld 1998; Bastiaanse & Thompson 2003; Buchert et al. 2008; Friederici 2017). In light of this evidence, it seems difficult to maintain a non-modular conception of the mind.

the maturation of language follows from the general maturation of the brain (Lenneberg 1967; Friederici 2017). At least two important states in language growth can be discerned: the initial state and the attained mature state (Chomsky 1976: 3). The initial state is the totality of the biologically predetermined (innate) language-related units and operations. The attained state is an I-language. As in the maturation of any other biological system, three factors play a role in the growth of language: innate factors (genetic endowment), experience (environment), and more general laws of nature (see Chomsky 2005 for details). Since different brains will be exposed to different linguistic experience during maturation, the attained states can be different (hence linguistic variation), but they can only vary within the range determined by the innate factors. An explicit characterization of the initial state of language growth is called *Universal Grammar* (UG). ‘Universal’ here refers to innate, biologically given cognitive capacities, not to surface-true generalizations about languages (understood non-technically) in the sense of Greenberg (1966) and related work. In other words, the main task of Universal Grammar is to precisely characterize the notion of a *possible human language*, without accounting for what is frequent, common, attested or currently attestable. An explicit characterization of the mature state of a language is called a *generative grammar*. Since phonological competence is an indispensable part of the mature state of language (see sections 2 and 3 for details), its description should be included in a generative grammar.

Because it is an aspect of the mind, an I-language is not directly observable.<sup>9</sup> Therefore, the main source of evidence for I-language is linguistic performance or the *use* of I-language (e.g., in speaking or understanding speech). Again, it is important to recognize clearly the difference between the *object of study* in linguistics (in all of its branches) and the *sources of evidence* for that study. The object of phonological study is

---

<sup>9</sup> The only sense in which an I-language could in principle be directly observed is in terms of brain activity. Since our current understanding of the neurobiological substrate of all modules of cognition, including I-language, is still very limited, the insight about the nature of I-language that can be gained on these grounds is also very limited.

phonological competence, which is a part of an I-language. The primary source of evidence bearing upon that object of study are spoken utterances. But spoken utterances, or speech more generally, are *not* the object of phonological study because by the time the human body outputs a spoken utterance, phonological competence has interacted with numerous factors external to linguistic competence (e.g., transduction procedures, rate of speaking, wearing dental braces etc.). Not discriminating between the object of study and evidence for that object leads to a serious confusion that hinders progress in both phonology and phonetics.

An example of this confusion can frequently be observed in discussions of gradience in phonology. The following is a fairly standard definition of ‘categoricity’ vs. ‘gradience’, and by emphasizing certain words in it, we wish to draw the reader’s attention to the conceptual level at which the definition is given:

[C]ategorical *sounds* (...) are stable and represent clear distinct phonological categories (e.g. *sounds* showing all characteristics of voiced segments *throughout* their *realizations*) (...); gradient *sounds* (...) may change *during* their *realization* and may simultaneously represent different phonological categories (e.g. *sounds* that *start* as voiced and *end* as voiceless). (Ernestus 2011: 2115)

From the emphasized words it is obvious that this characterization of categoricity and gradience is relevant to the domain of speech (performance), not grammar (competence). A problem arises when phonetic data are used to make inferences about phonology *directly*, as if every idiosyncratic datum recorded in speech or found in a corpus is relevant for phonology, without acknowledging the distance between competence and performance. Consider another passage from Ernestus (2011: 2118):

Ellis and Hardcastle (2002) found that four of their eight English speakers showed categorical place assimilation of /n/ to following velars in all tokens, two speakers showed

either no or categorical assimilation, and two speakers showed gradient assimilation. Together, the data show that place assimilation processes (...) may be gradient in nature. These processes cannot simply be accounted for by the categorical spreading of a phonological feature from one segment to another.

The cited results, showing inter- and intra-speaker variation, as well as both discrete and gradient effects, may constitute a salient illustration of the ubiquitous lack of uniformity in the *behavior* of members of a speech community, but it is not in the purview of phonology to provide an explanation of such phenomena. The fact that such variation “cannot simply be accounted for by the categorical spreading of a phonological feature from one segment to another” (*ibid.*), a claim most certainly true, does not mean there is something wrong with phonology conceived as categorical symbol manipulation. A plausible explanation of these results is that gradience is introduced into the data by the post-phonological transduction procedure that determines the temporal coordination of muscular activity (see section 4), and not by phonological computation itself. Notice the references to time highlighted in the above quote from Ernestus (2011: 2118), for example, “during” and “start as... end as”. Gradience involves change over time. If we think of phonology as involving a representational system (features and the like) and a computational system that can be regarded as a complex function of, say, composed rules (see sections 2 and 3), then there is no temporal aspect to phonology.<sup>10</sup> Questions about gradience in phonology are like questions about how fast a *wh*-element moves in syntax; both reflect a category error. The only kind of conclusion a phonologist can draw from the Ellis & Hardcastle experiment cited by Ernestus is that the I-language of some English speakers contains a feature-based rule that changes coronal nasals to velar nasals before velar consonants. If a featural assimilation rule correctly models a part of a speaker’s phonological competence, a phonetician can then posit hypotheses as to why such a pattern exists, why there

<sup>10</sup> Which is *not* to say that precedence relations, and (multi)linear ordering of phonological elements more generally, are irrelevant for phonology (see Idsardi & Raimy 2013).

is variability in externalization of this knowledge, what are the limits of its variation, whether the variation is purely biomechanical or partly/mostly/solely cognitive, and so on. Adopting such a perspective not only preserves a clear distinction between competence and performance, a necessity on many different grounds, but it also facilitates disentangling phonological conclusions from phonetic conclusions even though both are drawn from the same data.

## 1.2 CONSEQUENCES OF THE I-LANGUAGE PERSPECTIVE

Due to the nature of its object of study, generative linguistics assumes internalism, mentalism, realism, naturalism, and rationalism. These ‘-isms’ are philosophical stances that ultimately determine an overarching research method for the study of language that has proved fruitful.

As we have seen, *internalism* refers to the idea that the object of linguistic study is internal to the mind/brain, and that there is no coherent, scientifically useful notion of language external to the human mind/brain. Surely, there is no phonologist who thinks that language is completely outside of people’s brains. But just believing that there is something in the brain is not enough to make a linguist or a psychologist an internalist. An internalist claims that language is *only* in the mind/brain. As radical as this sounds, it is trivial to illustrate with a basic linguistic notion such as *word*. Inspection of the waveform of utterances demonstrates that there is no consistent correlation between the perceived boundary between words and the silent parts of the signal. The waveform of a spoken word like *stand* has silence between [s] and the release of [t], and a phrase like *There’s Mary* has no silence before the [m] of *Mary*. Howard Lasnik (2000: 3) has made this point forcefully:

The list of behaviors of which knowledge of language purportedly consists has to rely on notions like ‘utterance’ and ‘word’. But what is a word? What is an utterance? These notions are already quite abstract. Even more abstract is the notion ‘sentence’. Chomsky

has been and continues to be criticized for positing such abstract notions as transformations and structures, but the big leap is what everyone takes for granted. It's widely assumed that the big step is going from sentence to transformation, but this in fact isn't a significant leap. The big step is going from 'noise' to 'word'.

As Ray Jackendoff (1990: 164) puts it, "Language, as far as I can tell, is all construction." In the domain of phonology, the point has been made before: "It should be perfectly obvious by now that segments do not exist outside the human mind." (Hammarberg 1976). And even earlier, Edward Sapir (1933/1949: 46) recognized that the same principle holds for all domains of human cognition, stating that "no entity in human experience can be adequately defined as the mechanical sum or product of its physical properties." The internalist perspective thus maintains that the objects of cognition are better understood as our parses of the mind-external word, rather than as the mind-independent 'furniture' of that world (cf. Chomsky 1980: 218f). The related nativist perspective is that without the innate categories with which to parse the input, no learning can take place: "In any computational theory, 'learning' can consist only of creating novel combination of primitives already innately available" (Jackendoff 1990: 40, echoing Fodor).

Accordingly, linguistics is *mentalistic* in the sense that rather than being concerned with observable verbal behavior, it is concerned with 'mental aspects of the world' which stand alongside its mechanical, chemical, optical, and other aspects. The fact that the boundaries between mental and non-mental phenomena are not always completely clear is normal and unproblematic. As the eminent French philosopher Pierre Jacob (2010: 219) put it, "no philosopher of physical sciences believes that he is expected to offer a criterion for what constitutes mechanical, optical, electrical or chemical phenomena." Why, then, should mental phenomena be different?

*Realism* is just the reflection of the fact that language is a real aspect of the world: It is a certain subsystem of the human brain, and when the functioning of the brain is viewed from a sufficiently abstract point of view, language can be regarded as a mental



object. It commits us to admit that the apparent dualism between mind and brain is merely a result of our current inability to explain how an activity of neurons can give rise to thoughts, not a result of the belief that there indeed are two categorically different entities—the mind *is* the brain. This point was made by Vernon Mountcastle in a more general context of cognitive neuroscience:

The vast majority [of cognitive neuroscientists] believe in physical realism and in the general idea that no nonphysical agent in the universe controls or is controlled by brains. Things mental, indeed minds, are emergent properties of brains. Those emergences are not regarded as irreducible but are produced by principles that control the interactions between lower level events—principles we do not yet understand. (Mountcastle 1998: 1)

Since its object of study is a part of the natural world on a par with its mechanical, chemical, optical and other parts, linguistics can be regarded as a natural science. In his seminal book *Biological Foundations of Language* (1967), Eric Lenneberg, one of the founders of biology of language, noted that “we may regard language as a natural phenomenon—an aspect of our biological nature, to be studied in the same manner as, for instance, our anatomy” (p. vii). This *naturalism*, then, has certain methodological consequences, particularly in suggesting the validity of the use of standard research tools of the natural sciences such as simplification, isolation and idealization in linguistic inquiry. Thus, just like in physics, it is perfectly acceptable, and in fact necessary, to study language and its various subsystems in isolation and in a highly idealized setting in order to first gain an understanding of its fundamental nature. Only once that has been achieved to a reasonable degree can it be useful to introduce further complexity by studying the interaction of language with other systems. This view concerning the isolability of grammatical components follows that of philosophers of science such as Lawrence Sklar (2000: 54–55), who has even made the point that without such isolability, science itself would probably be impossible:

[W]ithout a sufficient degree of isolability of systems we could never arrive at any lawlike regularities for describing the world at all. For unless systems were sufficiently independent of one another in their behavior, the understanding of the evolution of even the smallest part of the universe would mean keeping track of the behavior of all of its constituents. It is hard to see how the means for prediction and explanation could ever be found in such a world [...] It can be argued that unless such idealization of isolability were sufficiently legitimate in a sufficiently dominant domain of cases, we could not have any science at all.

The standard *rationalist* stance is that the human mind contains innate knowledge—units and operations that exist in the mind independently of experience—and that this innate knowledge is a prerequisite for any kind of learning to take place: “the dimensions in terms of which experience is encoded cannot themselves be learned” (Jackendoff 2015: 187). We have already alluded to two arguments for nativism in phonology (see sections 2 and 3 for further discussion). First, there is the logical argument. Suppose that learning a given language involved learning that noun phrases like *dogs* come after adpositions like *with*. Such a rule can only be learned if the learner has access to the categories ‘noun phrase’ and ‘adposition’. Similarly, a phonological rule like ‘delete a consonant before another consonant (but not before a vowel or at the end of a word)’ can only be learned if the category ‘consonant’ can be applied to the data. Second, as suggested above, there is no guarantee that the categories of language have any consistent articulatory or acoustic correlates that enable those categories to be applied or triggered. For example, there are no obvious physical properties that correspond to the set of speech sounds [s, p, g, r, m, l] but not [i, u, a] and the end of a word: a *completely* unbiased scientist would not be able to derive the category ‘consonant’ from the signal. The degree of constriction in the vocal tract might be invoked as the relevant criterion here, but the cut-off point between the categories ‘consonant’ and ‘vowel’ is arbitrary and must be

predetermined for the parsing to begin. For a human speaker, the innate category is applied to a certain input, thus determining that consonants constitute a ‘natural class’. This problem of a *lack of invariance*—an absence of invariant properties that correlate with linguistic categories—is well known in the speech perception literature. Philosopher Irene Appelbaum (2006) suggests that there has been no progress in its solution despite half a century of work because the wrong question is being asked (see section 4).

The rationalist approach to the study of language is also based on the view that the data that we use to reason about language, e.g. the data that come from linguistic performance, represents the object of study in an *indirect* way. The reasons for this indirectness are clear: on the one hand, performance yields data that are necessarily intertwined with data about other cognitive systems and sensorimotor processes, and on the other hand, the data are qualitatively impoverished in comparison to the competence that produced them (in interaction with other systems). It is therefore not at all evident what constitutes ‘data’ for linguistics since (a) its object of study is not something directly observable, and (b) data do not come with a label that says that they are relevant for a particular object of study. For a phonologist qua cognitive scientist, it is necessary to understand that not all data from speech are relevant for phonology, and that it is therefore necessary to peel off the various complications that were introduced in the process of externalization from the underlying system of linguistic knowledge. In every scientific domain, it is the theory that determines whether something counts as relevant data or not. In other words, “we have to remember that what we observe is not nature in itself but nature exposed to our method of questioning” (Heisenberg 1962: 58). We have already seen the repercussions of this fact in the discussion of gradience in the preceding section. Thus, while linguistics *is* an empirical science, it gives epistemological precedence to theoretical constructs of great explanatory depth over the observables. Again, this is a perfectly normal state of affairs in the natural sciences. For example, the existence of black holes was maintained on purely theoretical grounds for at least a hundred years, before the first ever direct image of a black hole and its vicinity was published on April

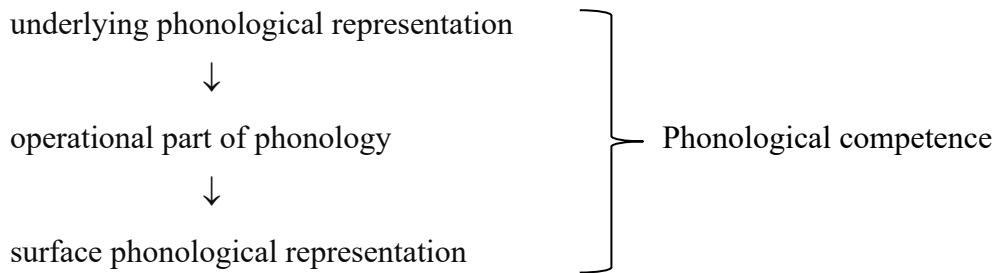
10<sup>th</sup>, 2019 (Akiyama et al. 2019).

To summarize, the object of study in phonology is phonological competence, a part of I-language and a module of the mind, which is to be studied by the classic, well-developed rationalist methods of natural science. Speech is merely one source of evidence for the study of that object, and it is always necessary to disentangle the phonologically relevant data from irrelevant artefacts of linguistic performance by giving priority to coherent theory over unfiltered observation.

## 2 THE THEORETICAL FRAMEWORK: SUBSTANCE FREE PHONOLOGY

The term ‘phonology’ is systematically ambiguous. On the one hand, it refers to a particular aspect of linguistic competence, and as such is equivalent to the term *phonological competence*. On the other hand, *phonology* is the name of the branch of linguistics that takes phonology in the first sense, i.e., phonological competence, as its main object of study. In the following discussion, this ambiguity is resolved by the context in which these terms are used.

Phonological competence is an aspect of mind that consists of elementary units and operations not completely shared by any other module of the mind, linguistic or otherwise. Phonological units are primitives from which the representational aspect of phonology is built, while the operations constitute the computational aspect of phonology. The conceptual minimum needed to describe phonological competence consists of two levels of representation and an operational part that connects these two levels, yielding a derivational scheme that is characteristic of most work in modern phonology:



In a fairly standard conception of the architecture of linguistic competence, the lexicon provides the syntactic component with minimal meaning-bearing linguistic units, morphemes, from which hierarchical structures are built. At the syntax-phonology interface, these hierarchical structures are flattened into linear strings of morphemes. A linearized string consisting only of the abstract forms of morphemes corresponds to the *underlying phonological representation (UR)*. The *surface phonological representation (SR)* is a data structure which the phonetic implementational system converts into movements of the speech organs. Since they are part of the same module, URs and SRs are constructed from the same type of units. These units can thus be said to comprise the *universal representational alphabet* of phonology from which all I-languages draw. The phonological representational alphabet consists of *phonological features*, and other distinctive, irreducible units of phonology such as *tones*, *moras* and *stresses*, i.e., elements of *prosody*. To constrain the exposition, we will primarily focus here on the nature of features, leaving the prosodic units aside.

The nature of phonological features can be elucidated by adopting the cognitive-neuroscience framework proposed by Gallistel & King (2010). There, what we have so far informally referred to as *units* are called *symbols*. Symbols are “physical entities in a physically realized representational system” (Gallistel & King 2010: 72), where the physical system in the case of phonological symbols, and all other cognitive symbols, is the human brain. Thus, phonological *features* are symbols realized in the human brain. The common properties of all neurobiological symbols are (at least) distinguishability,

combinability and 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. Of course, we are still far from being able to state how exactly features qua neural symbols are realized in the brain, but promising work such as Phillips et al. (2000), Scharinger et al. (2012), Hickok (2012), Bouchard et al. (2013), Mesgarani et al. (2014), Magrassi et al. (2015) and Scharinger et al. (2016) is emphasizing the importance of neural activity in the superior-most part of the superior temporal gyrus and sulcus, and Brodmann areas 44 and 6.

Features also meet the criterion of *combinability*: the hallmark of modern phonology is the notion that features can be grouped into sets to construct higher-level, non-atomic data structures. An unordered, unstructured set of features constitutes a phonological *segment* (see section 3 for details), while a particular organization of segments constitutes a data structure of the next higher taxonomic 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 the so-called *natural class behavior*—different structures can behave alike because they contain identical substructures.

Features are also an *efficacious* way of coding information since their combining leads to combinatoric explosion (Reiss 2012; Matamoros & Reiss 2016). 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 206 trillion different segments. With a number of the same order of magnitude, say fifty, the number of definable segments rises to  $3^{50}$  (about  $7 \times 10^{23}$ ). Of course, the richness that arises from feature combinability should not be taken to imply that any particular I-language should come close to exploiting the full range of possibilities. Instead, what we expect to find in particular I-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). The same point is made in the context of

genetics by biologist Peter Medawar (1960: 37):

The potential variation of human beings is enormously greater than their actual variation; to put it another way, the ratio of possible men to actual men is overwhelmingly large.

A corollary of this combinatoric explosion is that such richness goes a long way towards eliminating the need for a phonetics module specific to each language (e.g., Keating 1985; 1990), which simplifies the sequence of conceptual steps needed to account for the externalization of language.

Our outlook on the nature of phonological features is an elaboration of a long tradition in generative linguistics, where the general research agenda is to build biologically plausible cognitive models of language. It is therefore not surprising that Morris Halle already in the 1980s proposed that features be understood in terms of human neurobiology: “[T]he distinctive features correspond to controls in the central nervous system which are connected in specific ways to the human motor and auditory systems” (Halle 1983/2002: 109). Even though some progress has been made in discovering the exact nature of Halle’s proposal (e.g. Mesgarani et al. 2014), we are still far from being able to refer to features by stating their neurobiological substrate and therefore have to resort to using symbols 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. To reiterate, 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) interpret the identity of a feature by the place of the neural activity (or a combination of the activity’s place and firing rate) in the brain. 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

2010: 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/or the rate of its repetition is how the transducer determines the identity of the feature and ‘knows’ which neuromuscular schema (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. 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.

Features are manipulated by phonological rules, which should be understood as neural functions in the sense of Gallistel & King (2010: §3). The totality of the primitive phonological symbols and functions that manipulate them constitute the phonological module of the language faculty. Also, features are interpreted by particular transduction algorithms at the interface between the phonological module and the sensorimotor system. The Cognitive Phonetics theory (Volenec & Reiss 2017; 2019), described in section 4, spells out how this transduction proceeds and provides hypotheses about how it is realized neurobiologically.

The symbolic nature of something 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*, *correlate*, or *the represented*. Phonological features are symbols which refer to aspects of speech. At this point, it is of utmost importance not to “make the common mistake of confusing the symbol with what it represents” (Gallistel & King 2010: 56); “the tendency to confuse symbols with the things they refer to is so pervasive that it must be continually cautioned against” (p. 62). There is a connection between phonological features and speech, but this connection is highly complex and indirect (see section 4), and features *do not* encode speech-related information in any straightforward way. This was the crucial characteristic of Distinctive Feature Theory from its inception:



Such considerations [that languages do not make free use of acoustic values or articulatory properties] were much in our minds thirty years ago 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: 108–109)

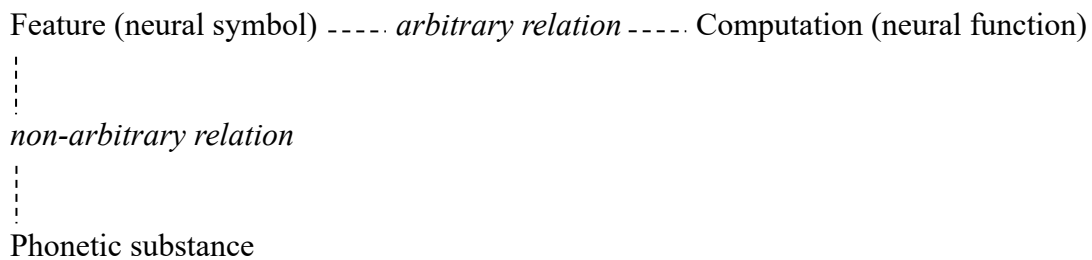
In linguistics, information related to speech is called *phonetic substance*. It is the totality of the articulatory, acoustic and auditory properties and processes that constitute speech. For example, properties and processes of speech such as movements of the tongue, values of formants, loudness, duration 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 is just an instance of the aforementioned mistake of confusing the symbol with what it represents.

The relation between any given feature and its correlate is not random or arbitrary. If it were, then any feature could in principle be realized by any possible human articulation, similarly to how the concept/signified ‘DOG’ can in principle be assigned to any possible signifier. If such Saussurean arbitrariness were applicable to the realization of features, then it should be possible that +ROUND in one language gets realized as a lowering of the velum, and in another as a raising of the tongue dorsum, and so on. Instead, there is a non-arbitrary, lawful relation between features and their correlates. The nature of this relation is described in more detail in section 4, but at this point it is important to emphasize that the lawful relation between features and their correlates is *phonologically* irrelevant. That means that phonological computation treats features as invariant

categories, manipulating them in an arbitrary way irrespective of the variability in their realization in speech and irrespective of phonetic substance in general.

This fact can be clearly observed in the example of Turkish vowel harmony, where certain suffixes will contain different vowels depending on the preceding root vowels. The [+ROUND] root vowels [u] and [o] give rise to a suffix form with [u], whereas the corresponding [−ROUND] vowels [ʊ] and [ɑ] give rise to a suffix form with [ʊ]. The phonetic correlate of [+ROUND] for [u] is systematically different from the phonetic correlate of [+ROUND] for [o]. Because the lower jaw is lowered to a greater extent in the articulation of [o] than in the articulation of [u], the lip rounding for [o] will always take a different shape than the lip rounding for [u]. Crucially, the operation of vowel harmony will always ignore this obvious phonetic difference, and will treat [+ROUND] as an invariant category—by virtue of containing the substance-free invariant category [+ROUND], both [u] and [o] will trigger the operation of vowel harmony in the same way, giving rise to a suffix with [u]. The relation between features and phonological computation is thus completely independent from the relation between features and phonetic substance, as can be seen in this diagram:



The diagram summarizes the fundamental tenet of *Substance Free Phonology*: Even though there is a lawful relation between features and phonetic substance, both the representational and the computational aspect of phonology are devoid of phonetic substance (cf. Hale & Reiss 2008).

We maintain, as the null hypothesis, the view that Universal Grammar provides a

finite set of discrete phonological features. This position currently appears to be out of favor among many phonologists, although anyone who accepts the innateness of constraints that refer to specific features (e.g., IDENT-IO<sub>[VOICED]</sub>) in various versions of Optimality Theory (OT), including the original proposals of Prince & Smolensky (1993/2004) and McCarthy & Prince (1995), must be committed to the innateness of features.

Work that denies the existence of innate features suffers from at least two failings. First, as a whole, this literature (e.g., Carr 2006, Mielke 2008, Archangeli & Pulleyblank 2015) ignores a logical argument presented in various forms, at least since Kant concluded in the eighteenth century that “all objects of experience must necessarily conform” to cognition:

Up to now it has been assumed that all our cognition must conform to the objects; but all attempts to find out something about them a priori through concepts that would extend our cognition have, on this presupposition, come to nothing. Hence, let us once try whether we do not get farther with the problems of metaphysics by assuming that the objects must conform to our cognition, which would agree better with the requested possibility of an a priori cognition of them, which is to establish something about objects before they are given to us. (Kant, cited in Rohlf 2016)

This conclusion has been reiterated and explicated over decades in the generative literature (e.g., Fodor 1980; Hale & Reiss 2003). For example, Jackendoff (2015: 187) points out that:

If one is studying item-based or exemplar-based learning—if one thinks that learning consists of encoding all the instances of whatever a person has experienced in some domain, and that understanding novel inputs is a matter of analogy or interpolation among instances [references deleted]—then one ought to be asking: Exactly how does the mind/brain encode instances it has encountered, and what are the dimensions available

for encoding them? A crucial constraint on such a theory is that the dimensions in terms of which experience is encoded cannot themselves be learned. They form the basis for learning; they are what enable learning to take place at all. So they have to somehow be wired into the brain in advance of learning.

Until this logical argument for the innateness of features is addressed, so-called ‘emergentist’ theories cannot be taken seriously. It is noteworthy that even some researchers in the AI machine learning domain have recently recognized the need to build ‘priors’ into their systems (Versace et al. 2018).

A second, oft-repeated argument against innate features is the putative existence of rules that do not refer to sets of segments that constitute a natural class (Mielke 2008, *inter alia*). However, work that denies an innate feature set in this way does not provide a general and explicit theory of rules, so such claims cannot even be evaluated. In Logical Phonology (see section 3), rules refer to natural classes by definition: a statement that cannot be formulated in terms of natural classes is not a rule. An example that can clarify that point can be drawn from Halle’s paper *Knowledge Untaught and Unlearned* (1978/2002). Halle explores the productivity of English plural formation and essentially presents a phonological Argument from the Poverty of the Stimulus (PoS), although he does not use that term explicitly. His main point is that rules refer to natural classes defined by shared features, and not to sets of atomic segments. The argument relies on predicting that an English speaker exposed to the pronunciation of *Bach* as [bax], with a voiceless velar fricative, can form the plural [baxs] despite never having encountered the devoicing of underlying /z/ after /x/. The implicit reasoning is that a speaker has learned a rule that devoices /z/ after /p, t, k, θ, f/, and that /x/ falls into any natural class that contains those segments. So, the pronunciation [baxs] is just normal rule application, if one thinks at the level of features. In other words, the rule refers to a natural class that is defined *intensionally*, but the *extension* of the set of relevant trigger segments changes when [x] is introduced. It turns out that Halle’s PoS argument from the plural can be

pushed further, and the relevance of the intensional characterization of rules can be made even more apparent.

The English plural form [-iz] occurs in a natural class of environments, as expected, because the vowel insertion is triggered ‘between coronal stridents’. The form [-z] does *not* occur in a natural class of environments, because it is the ‘elsewhere’ case, identical to the underlying form. However, the form [-s] arises by a devoicing rule, but it does *not* occur in a natural class of environments *on the surface*—the voiceless coronal stridents [s, ʃ, tʃ] are in the natural class that contains [p, t, k, f, θ] (assuming [f] is +STRIDENT). The well-known solution is that the vowel insertion rule precedes and bleeds the devoicing rule. Intensionally stated, devoicing applies to a natural class of segments (all the voiceless obstruents), but the data available to the learner contradicts this. We therefore argue for an interpretation of PoS that not only *allows* the learner to posit rules in terms of natural classes despite contradictory surface patterns, but actually *forces* them to do so.

Much of the literature says that rules ‘typically’ refer to natural classes, and as noted above Mielke (2008) proposes that the existence of rules that do not rely on natural classes constitutes an argument against an innate feature set. In contrast, we propose that natural classes are defining properties of rules: *If a statement cannot be made in terms of natural classes, then that statement is not a rule.* We argue that it is only possible for linguists to study the range of possible rules if one stipulates that rules are based on natural classes. By defining rules in terms of natural classes, one severely restricts the hypothesis space for an acquirer—only processes definable in terms of natural classes can be part of the target phonological grammar. As pointed out above, the [-s] form of the plural does not occur in a natural class of environments, but the devoicing rule intensionally applies in a natural class of environments. If UG determines that rules must be stated in terms of natural classes, then a learner formulates such rules even in light of contradictory evidence—this is a PoS argument.

The view of natural classes developed here parallels the phenomenon of *amodal completion* (Michotte & Burke 1951) in visual perception and other ‘masking’ effects. *Figure 1* is interpreted by the visual system as two objects, a magenta one partially occluding a blue one, despite the lack of continuity between the two blue regions. The visual system infers the part of the blue rectangle missing from the signal. Similarly, in acquiring ‘English’, the language acquisition

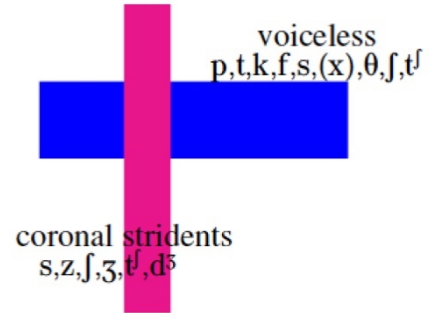


Fig.1: Amodal completion

device infers a single voiceless obstruent class as a trigger of devoicing, even in the absence of direct evidence. As the IPA symbols in *Figure 1* show, we can think of the magenta region as representing the natural class of coronal stridents, which trigger vowel insertion. Because of the bleeding order, this rule ‘occludes’ some members (s, ʃ, tʃ) of the intensionally defined natural class which triggers devoicing, the set of all voiceless obstruents. Just as our visual systems infer a single blue object, our phonological acquisition systems infer a single voiceless obstruent natural class. If learners could make a separate rule for each segment that devoices /z/, then Halle’s (1978/2002) claim about devoicing after /x/ would not be justified. We propose that even in the face of contradictory evidence (the partial ‘occlusion’ by coronal stridents of the natural class of voiceless obstruents), a learner constructs a single rule for devoicing. Without this insight—the centrality of feature-based natural classes, despite the poverty of the stimulus—it is impossible to develop a universal theory of rule formulation and rule interaction.<sup>11</sup>

In light of these logical and methodological considerations, we maintain that UG

<sup>11</sup> The masking effect we just saw in vision is paralleled in auditory perception (Bregman 1990). If our parallel to phonological learning is on the right track, we can take this as further corroboration of the cognitive science approach to phonology advocated in this paper, since the notion of masking seems relevant to these three domains. Conversely, we might think of amodal completion in vision and audition as versions of the PoS argument, since they involve inference of structures for which there is no direct evidence in the signal.

provides a set of innate features. Since there are sound logical and empirical reasons to doubt that features can be learned, their innateness is the best null hypothesis. Numerous legitimate challenges remain for the idea that features are innate, including the question of whether sign languages can plausibly be encoded with the same set of features: Is a bilingual native speaker/signer of English and ASL using the same set of features for both languages, ones that we just happen to refer to by names like ‘ROUND’ and ‘VOICED’? How many features are there? Is there any reason to think that the twenty or so features assumed in much of the literature are even close to sufficient?<sup>12</sup>

We mentioned in section 1.1 that there is a logical relationship between internalism and nativism: a denial of internalism entails a denial of nativism, since nativism is a claim about innate *internal* knowledge. We have argued that phonology should be treated like syntax with respect to issues like internalism and nativism. The following are some statements showing that this view is by no means universally held in the field:

[H]owever much poverty of the stimulus exists for language in general, there is none of it in the domain of the structure of words, the unit of communication I am most concerned with. Infants hear all the words they expect to produce. Thus, the main proving ground for UG does not include phonology. (MacNeilage 2008: 41)

Peter MacNeilage is not a phonologist, but rather a psycholinguist who works on speech, so it is not surprising that his perspective is far from that of the orthodoxy of generative

---

<sup>12</sup> Gunnar Fant (1969) considered the feature system presented in the *SPE* to be “an improvement over that of Jakobson, Fant & Halle (1952)”. However, he pointed out that

there is a danger that the impact of the theoretical frame with its apparent merits of operational efficiency will give some readers the impression that the set of features is once for all established and that their phonetic basis has been thoroughly investigated. This is not so.

Of course, Fant’s conclusion still holds. Research on features is not complete, any more than research on genes is complete. There is no reason to expect any contemporary phonology textbook or article to list the correct set of features. Even the validity of appeals to underspecification—the assumption that features can be valued ‘+’, ‘−’ or be absent from a segmental representation—is open for discussion.

linguistics, especially syntax. Jeff Mielke is a linguist by training, and the clarity with which he rejects the generative tradition is refreshing:

Many of the arguments for UG in other domains do not hold for phonology. For example, there is little evidence of a learnability problem in phonology (see Blevins 2004 for discussion). (Mielke 2008: 33)

Most of the evidence for UG is not related to phonology, and phonology has more of a guilt-by-association status with respect to innateness. (Mielke 2008: 34)

Archangeli and Pulleyblank are also quite clear in rejecting the generative tradition and the idea that phonology is like syntax, not least by the title of their paper *Phonology without universal grammar*:

See Mielke (2004) on why features cannot be innately defined, but must be learned. (Archangeli & Pulleyblank 2015: 2)

[Children face] the challenge of isolating specific sounds from the sound stream. (*ibid.*)

[T]he predictions of Emergent Grammar fit the data better than do the predictions of UG (*op. cit.*: 4)

In the first quotation, Archangeli and Pulleyblank reject innate features. In the second one they suggest, contrary to our discussion of Jackendoff, Sapir and Hammarberg, that segments exist in the sound stream. In the third one they claim that UG is empirically a worse hypothesis than their alternative Emergent Grammar. Phonological PoS and the innateness of features is also rejected by Philip Carr:

Phonological objects and relations are internalisable: there is no poverty of the stimulus



argument in phonology. No phonological knowledge is given by UG. (Carr 2006: 654)

Finally, Juliette Blevins rejects PoS by suggesting that the auditory input to the learner is sufficient to capture all relevant generalizations:

Within the domain of sounds, there is no poverty of the stimulus. (Blevins 2004: 235)

[I offer] general arguments against the ‘poverty of stimulus’ in phonology. [...] To my knowledge, there is no argument in the literature that phoneme inventories, stress patterns, tone patterns, phonotactics, and regular phonological alternations cannot be acquired on the basis of generalizations gleaned directly from auditory input. (*op. cit.*: 66)

All of these scholars seem to be ignoring standard views about what it takes to develop an acquisition theory:

[T]he development of the theory of grammar [UG], and intensive application of this theory, is a necessary prerequisite to any serious study of the problem of language acquisition and many other problems of immediate psychological significance. This theory must contain a precise specification of the class of formalized grammars; it must provide a precise definition of this notion, that is, an explicit schema and notation for grammars and a specification of the conditions that a set of rules, stated in these terms, must meet in order to qualify as a grammar of some natural language. (Chomsky 1962: 534)

Without an explicit theory of phonological grammar, Blevins, Mielke and the others cannot, by fiat, declare that there is no learnability problem for phonology. Contrary to Blevins’ (2004: 66) claim that “there is no argument in the literature” that aspects of phonology “cannot be acquired on the basis of generalizations gleaned directly from auditory input”, what we find in various forms throughout the history of modern phonology, for example in the work of Sapir, Jakobson, Halle, Hammarberg and others, is that exact

argument, as we have shown in the preceding sections.

It is of interest that Archangeli & Pulleyblank and Blevins cite approvingly the work of Terence Deacon:

I think Chomsky and his followers have articulated a central conundrum about language learning, but they offer an answer that inverts cause and effect. They assert that the source of prior support for language acquisition must originate from *inside* the brain, on the unstated assumption that there is no other possible source. But there is another alternative: that the extra support for language learning is vested neither in the brain of the child nor in the brains of parents or teachers, but outside brains, in language itself. (Deacon 1997: 105)

Over countless generations languages should have become better and better adapted to people,  
so that people need to make minimal adjustments to adapt to them. (*op. cit.*: 107)

Grammatical universals exist, but I want to suggest that their existence does not imply that they are prefigured in the brain like frozen evolutionary accidents. In fact, I suspect that *universal rules or implicit axioms of grammar aren't really stored or located anywhere*, and in an important sense, they are not determined at all. (*op. cit.*: 115–116; emphasis added)

Of course, none of this makes sense from the I-language perspective. Despite the fact that Deacon (1997) studies the evolution of the brain and discusses, for example, the “loss of language abilities due to brain damage (termed *aphasia*)” (p. 280), he seems to locate language out in the world, not in the head. He rejects internalism. If we assume that the fact that Archangeli & Pulleyblank and Blevins cite Deacon can be read as an acceptance of his rejection of internalism, we can now see why they are driven to reject PoS and nativism. Without internalism, there is no nativism.

To reiterate how our proposals differ from these, let us summarize some of the lessons of this section. Words, segments and features are not ‘in’ the acoustic signal. These are categories that our minds impose on the signal. Without an innate inventory of such categories there can be no learning. Even granting the existence of such categories, the surface patterns of language do not reflect in a straightforward manner the nature of phonological computation, since even natural class behavior of segments can be ‘occluded’ by factors like rule-ordering, as we see in the English plural. These considerations support the idea that “a necessary prerequisite to any serious study of the problem of language acquisition [is] a precise specification of the class of formalized grammars” (Chomsky 1962: 534). In other words, we cannot even start to explain acquisition without an explicit theory of rules and their interaction. In the next section, we will begin to flesh out a model of rules built on operations that apply to formally specified natural classes of segments. Phonetic substance will play no part in this formal model.

### 3 A WORKING MODEL FOR PHONOLOGY: LOGICAL PHONOLOGY

Much of the phonological literature since *The Sound Pattern of English* has been concerned with accounting for *markedness*, inspired by the possibility that the model developed in the first eight chapters of the book was “overly formal”:

The problem is that our approach to features, to rules and to evaluation has been overly formal. Suppose, for example, that we were systematically to interchange features or to replace  $[\alpha F]$  by  $[-\alpha F]$  (where  $\alpha = +$ , and  $F$  is a feature) throughout our description of English structure. There is nothing in our account of linguistic theory to indicate that the result would be the description of a system that violates certain principles governing human languages. To the extent that this is true, we have failed to formulate the principles of linguistic theory, of universal grammar, in a satisfactory manner. In particular, we have not made use of the fact that the features have intrinsic content. (Chomsky & Halle 1968:

400)

For example, work in classical OT, which relies on two classes of constraints, one of which is known as the *markedness* or *well-formedness constraints*, cautions against extreme formalism:

We urge a reassessment of [the] essentially formalist position. If phonology is separated from the principles of well-formedness (the “laws”) that drive it, the resulting loss of constraint and theoretical depth will mark a major defeat for the enterprise. (Prince & Smolensky 1993/2004: 234)

Rather than following up on the initial worries expressed at the beginning of *Chapter 9* of the *SPE*, Logical Phonology maintains the ultimately strict formalist position, expressed later in the same chapter:

It does not seem likely that an elaboration of the theory along the lines just reviewed will allow us to dispense with phonological processes that change features fairly freely. The second stage of the Velar Softening Rule of English (40) and of the Second Velar Palatalization of Slavic strongly suggests that the phonological component requires wide latitude in the freedom to change features, along the lines of the rules discussed in the body of this book. (Chomsky & Halle 1968: 428)

In other words, phonological theory should define a combinatoric space defined by representational and computational primitives. The characterization of which particular regions of this phonological combinatoric space are actually or potentially attested, or are more or less common, is within the purview not only of the diachronic linguist, the phonetician and the acquisitionist, but also of the historian and archaeologist who studies patterns of migration, settlement, conquest and genocide that have produced the current distribution of language families in the world. For the construction of a general

mentalistic phonological theory, the frequency of occurrence of a particular phonological phenomenon is irrelevant, since that frequency is determined by political, social, economic, geographical and other extra-linguistic factors. As we have seen in section 2, the combinatoric space provided even by a modest set of features is exceedingly rich, and there is plenty of work for phonologists to do to further elaborate a formal model of computation and representation.

*Logical Phonology* (LP) maintains the SPE view of the phonology of a language as a complex function that maps input representations to output representations. This function is the composition (in a specific order) of more basic functions, typically called the *rules* of the language. As a first approximation, each rule maps strings of phonological segments to other strings of segments. As Gallistel & King (2010) point out, functions are operations that take variables as input, and this property applies to phonological rules, too. For example, a traditionally expressed rule like ‘ $t \rightarrow s / \_ i$ ’ maps members of the class of strings such that *ati* maps to *asi*, *okutilo* maps to *okusilo*, *aka* maps to *aka* and *mokitapi* maps to *mokitapi*. These mappings are not arbitrary—the form of the input determines the form of the output. In this simple example, rules apply not to atomic segments, but to strings, which are a more complex data structure. Phonologists know that even segments have an internal feature composition, and that a more complete model of the nature of rules available for phonological computation requires an elaboration of the taxonomy of phonological symbols. We need to get to the point where we can formalize rules such as ‘coronal stops become their corresponding fricative before any high front vowel’.

Symbols in computation must be physically realized entities that are manipulated by the programming and architecture of a computational system. Because the brain is physically limited in power and space, and because scientific reasoning leads us to propose the simplest possible hypothesis, a desirable model is one with a small number of primitive symbols and basic processes for combining them. By combining primitive symbols, more complex symbols can be created when new stimuli require new

representations. Gallistel & King (2010: 79–80) themselves use a linguistic example to illustrate the idea of a basic symbol inventory in both engineered and natural systems:

No language in the world has a word for the message, “After the circus, I’m going to the store to get a quart of skim milk.” (...) Minimizing the number of atomic data is desirable in any symbol system as it reduces the complexity of the machinery required to distinguish these data. This is why computers use just two atomic data and likely why nucleotide sequences use only four.

The atomic data enter into combinations that lead to more complex symbols, according to the following taxonomy:

- (1) Gallistel & King’s (2010: 79) symbol taxonomy
  - atomic data
  - data strings
  - nominal symbols
  - encoding symbols
  - data structures

For example, atomic data can be ordered in strings, and strings can be arranged into complex data structures like trees with precedence and containment relations among strings. Logical Phonology adopts a model of representation consistent with the symbol taxonomy in (1).

Vowels and consonants are *segments*. That is, they exist at the level of organization of language that we commonly represent with the symbols of the International Phonetic Alphabet. However, each segment symbol in LP is actually an abbreviation for a *set*

of valued features.<sup>13</sup> These sets are *consistent*, which means that they cannot contain incompatible values;<sup>14</sup> for a given feature F, a segment cannot contain both –F and +F. We can thus think of the vowels in the words *beet* and *boot* as sets of valued features, as shown in (2).

(2) Specification of sets of valued features corresponding to the vowels /i/ and /u/

$$/i/ = \{-\text{BACK}, -\text{ROUND}, +\text{HIGH}\}$$

$$/u/ = \{+\text{ROUND}, +\text{HIGH}, +\text{BACK}\}$$

These sets of features are part of what Poeppel (2012: 35) calls “the infrastructure of linguistics [that] provides a body of concepts that permit linguists and psychologists to make a wide range of precise generalizations about knowledge of language”. Realism (see section 1.2) commits us to accept that these sets of features are also part of language itself, not just tools that linguists use to describe language: “Since linguistics is an empirical science, the only justification for developing a theoretical framework is that the framework mirrors relationships among the facts of natural languages” (Halle 1975: 532–533; see also Hammarberg 1976: 355). *Table 1* illustrates more fully the data structures relevant to phonology.

<sup>13</sup> “The phonetic symbols [p], [t], [e], [i], [u], etc., are simply informal abbreviations for certain feature complexes; each such symbol, then, stands for a column of a matrix of the sort just described.” (*SPE*, p. 5)

<sup>14</sup> For the sake of simplicity, we are abstracting away from the so-called *contour segments*, such as affricates and diphthongs.

**TABLE 1.** A HIERARCHICAL TAXONOMY IN LANGUAGE.

Level	Notation	Name
Atomic	$\{+, -\}, \{\text{HIGH, BACK, ROUND, etc.}\}$	coefficients and features
Ordered pairs	$\langle +\text{ROUND} \rangle$	(valued) features
Sets of ordered pairs	$\{ \langle +\text{ROUND} \rangle, \langle -\text{HIGH} \rangle, \langle +\text{BACK} \rangle \}$ abbreviated [k], [i], [θ], [p] etc.	segments
Strings of sets of ordered pairs	/kæt/	strings (phonology of morphemes)
Tripartite structures	/kæt/ w/ syntactic & semantic info	morphemes
Lists of morphemes	(a complex structured list)	mental lexicons

The atomic value and feature symbols are combined into ordered pairs in the process of parsing speech. Such ordered pairs are grouped into sets corresponding to segments. Minimal units of meaning, called *morphemes*, are stored with information related to meaning and syntactic category (VERB vs. NOUN, for example). Of course, the segments that constitute the phonological representation of a morpheme must be stored in a specific order since the same three segments [æ, k, t] can be ordered to express more than one meaning, as in *act*, *tack*, *cat*.

Given this infrastructure for symbols, we can start to formalize the treatment of these symbols by a computational system. LP defines *possible rules* as functions that map sequences characterized by natural classes of segments to other sequences. Since each segment is treated as a set of valued features (which in turn are ordered pairs), we have available a simple method of formalizing the sometimes intuitively stated notion *natural class*:

**Definition:** The smallest natural class containing a set of segments  $S = \{s_1 \dots s_n\}$  is defined by the generalized intersection of S. The generalized intersection of a set of sets  $R = \{r_1 \dots r_m\}$  is just the set of elements that are in every set  $r_i$ .



This can be illustrated using the vowels /i/ and /u/ specified in (2). The generalized intersection  $\cap\{i, u\} = \cap\{\{-\text{BACK}, -\text{ROUND}, +\text{HIGH}\}, \{+\text{ROUND}, +\text{HIGH}, +\text{BACK}\}\} = \{+\text{HIGH}\}$ . The segments in the smallest natural class containing /i/ and /u/ are thus all those that are supersets of the set  $\{+\text{HIGH}\}$ . This natural class is described intensionally using standard set theory formalism thus:  $\{x : x \supseteq \{+\text{HIGH}\}\}$ . For ease of exposition, we will denote such a natural class using the traditional square brackets for the classes that constitute targets and environments of phonological rules, thus  $[+\text{HIGH}]$ . It follows, then, that if a language has a rule that refers to the vowels /i/ and /u/, then the rule must apply equally to the vowel /y/ =  $\{+\text{ROUND}, +\text{HIGH}, -\text{BACK}\}$ , if this vowel also occurs in the language, because /y/ is also a superset of the generalized intersection  $\cap\{i, u\}$ . This result formalizes the notion of *generalizing beyond the data*, although viewed at the correct level of abstraction, there is no generalization, just application of the rule to whatever value has been assigned to that variable serving as its input symbol. It also follows that if in some language a putative process affects /i/ and /u/ but not /y/, then that process cannot correspond to a single phonological rule. Adopting a simple, but explicit model like this allows us to leverage the well-established results of set theory, and also leads us to understand apparent exceptions as arising from, say, the masking effects of rule ordering, as in the discussion of the English plural in section 2. In other words, by having a theory of the scope of rules, we have a theory of the limits of rules—certain patterns that might be considered the result of a single rule in an informal analysis, have to be accounted for with two rules, when the theory progresses and becomes more explicit.

The approach we describe is an example of more general principles that have been advocated in linguistics for decades:

By pushing a precise but inadequate formulation to an unacceptable conclusion, we can often expose the exact source of this inadequacy and, consequently, gain a deeper understanding of the linguistic data. More positively, a formalized theory may automatically provide solutions for many problems other than those for which it was explicitly

designed. Obscure and intuition-bound notions can neither lead to absurd conclusions nor provide new and correct ones, and hence they fail to be useful in two important respects. (Chomsky 1957: 5)

Note that Mielke's (2008) method for rejecting natural class reasoning is to scan the literature for 'rules' that do not apply to natural classes. One problem is that he counts as a rule anything that has been called a rule by his sources. In the absence of a formal theory of rules, just about any descriptive statement can be regarded a rule; of course, science cannot be done in such loose terms. Another problem is that this technique should come up with the English devoicing rule as an example that does not use natural classes, since on the surface, it is not triggered by a natural class of segments (see section 2). In the case of English, it is obvious how to save a natural class solution through a bleeding rule relationship, but for the many less-studied languages he cites, from Abujmaria to Zina Kotoko, it is unlikely that his sources are as well-documented and sophisticated as what we have for English.

Now that we have a better understanding of natural classes, we turn to considering another aspect of phonological computation, the nature of rules. Following Bale et al. (2014), we suggest that it is necessary to deconstruct the traditional arrow of phonological rules, because there are in fact several distinct operations at work in phonological computation. Here we focus on one such operation, *priority union*, which, as we show, allows a unified analysis of feature changing and feature filling processes.

Regular set union combines the elements of two sets into a set whose members are all the elements that are in at least one of the input sets. So, given sets A and B,  $A \cup B = C$ , where each element of C is in either A or B or both:  $\{w, x, y\} \cup \{w, y, z\} = \{w, x, y, z\}$ .

Applying the union to a segment (which is a set of valued features) and another set of valued features potentially leads to an inconsistent output (3).

(3) Union can yield an inconsistent output

$$\{+COR, -SON, +VOICED\} \cup \{-VOICED\} = \{+COR, -SON, +VOICED, -VOICED\}$$

Because the output contains both +VOICED and –VOICED as members, it is inconsistent. In order to maintain the requirement that segments are consistent, an operation different from simple union is needed for phonological rules. Adapting a definition from Kaplan (1987), we make use of *priority union*, which is defined in (4). In this definition, the valued features ‘+F’ and ‘–F’ are referred to as *opposites*.

(4) Definition of priority union

The priority union of two sets of valued features, A and B, is the union of the set of valued features that are in A, and the set of valued features in B whose opposites are not in A.

Unlike simple set union, priority union is not commutative—the priority union of A with B is not necessarily the same as the priority union of B with A. Paraphrasing Dahl (2002), priority union (also called *default unification*) takes two feature sets, one of which is identified as *strict*, while the other one *defeasible*, and combines the information in both such that the information in the strict set takes priority over that in the defeasible set. We denote priority union by the normal union symbol with an arrow underneath: ‘ $\cup\rightarrow$ ’ or ‘ $\cup\leftarrow$ ’ (for reasons that will become apparent below). The arrow points from the strict argument to the defeasible argument, as shown in (5).

(5) Priority union examples

- a.  $\{+F\} \cup \rightarrow \{+G\} = \{+F, +G\}$
- b.  $\{+F, -G\} \cup \rightarrow \{+G\} = \{+F, -G\}$
- c.  $\{+F, -G\} \cup \rightarrow \{+G, +H\} = \{+F, -G, +H\}$
- d.  $\{+F, -G\} \cup \leftarrow \{+G, +H\} = \{+F, +G, +H\}$

In (5a), priority union gives the same result as normal union. In (5b), the valued feature  $-G$  of the strict argument takes priority over the  $+G$  of the defeasible argument. The same situation holds in (5c), except that the defeasible argument contributes  $+H$ , since there is no conflict with the strict argument. In (5d), the arrow has been reversed, so that  $+G$  is a member of the strict argument and  $-G$  is in the defeasible argument.

Inkelas (1995) analyses Turkish as having a three-way underlying consonant contrast that can be used to illustrate how priority union can capture feature-filling phonological rules. The segment  $/d/$  is a voiced coronal stop,  $/t/$  is a voiceless coronal stop, and  $/D/$  is a coronal stop with no specification for voicing.<sup>15</sup> The fully specified stops do not alternate in the relevant environments:  $/t/$  and  $/d/$  surface unchanged in onsets and codas, but  $/D/$  surfaces as  $[d]$  in onsets and as  $[t]$  in codas.

Given the set-theoretic formalization, there is no way to define a natural class that (intensionally) targets just  $/D/$ , but there is a way to make a rule whose effect *appears* to target just  $/D/$  by using priority union. Rule (6) turns  $/D/$  to  $[t]$  in codas.

<sup>15</sup> The exact analysis of Turkish is controversial, but there exist other phenomena that share the same structure as the presentation here, so we can abstract away from phonetic details—we are just concerned with three underlying segments  $a$ ,  $b$ ,  $c$ , such that  $b$  is the intersection of  $a$  and  $c$ . In other words,  $b$  is unspecified for a feature with respect to which  $a$  and  $c$  have opposite values. Allowing underspecification makes for a more ‘stripped-down’ theory, because it does away with the stipulation that segments must be complete, i.e., that they must have a value for each feature. Removing the stipulation of completeness is parallel to removing the stipulation against Internal Merge in syntax, and thus unifying Merge and Move (Chomsky 2007). Additionally, underspecification gives the model more descriptive power for a given number of features, since the number of definable segments with  $n$  features in UG then increases from  $2^n$  to  $3^n$ . This increase in descriptive power has at least two benefits. First, it means that it is easier to resist the temptation of positing a language-specific phonetic module, as discussed below in section 4. Second, it is not necessary to posit as many features in UG to achieve a given level of descriptive power. Obviously, it is desirable to reduce the amount of information that we need to attribute to the genetic basis of language.

- (6)  $[+COR, -CONT, -SON] \cup \{-VOICED\}$  / in Coda position

Recall that this target specification refers to all segments that are supersets of the set of valued features  $\{+COR, -CONT, -SON\}$ , so the target natural class contains the three segments /t, D, d/. The target and the individual segments are defined as in (7).

- (7) A natural class with an underspecified member<sup>16</sup>

Description of class:

$$[+COR, -CONT, -SON...] = \{x : x \supseteq \{+COR, -CONT, -SON...\}$$

Members of the class:

- a. /t/ =  $\{-VOICED, +COR, -CONT, -SON...\}$
- b. /d/ =  $\{+VOICED, +COR, -CONT, -SON...\}$
- c. /D/ =  $\{+COR, -CONT, -SON... \text{ (no VOICED feature)}\}$

Priority union of each of the Turkish stops with the set  $\{-VOICED\}$  is shown in (8).

- (8) Priority union for Turkish codas

- a.  $/t/ \cup \{-VOICED\} = [t]$
- b.  $/D/ \cup \{-VOICED\} = [t]$
- c.  $/d/ \cup \{-VOICED\} = [d]$

<sup>16</sup> The ellipsis ‘...’ denotes whatever other features may be relevant to characterizing the segments in question, in this case anything except VOICED.

In (8a), priority union is applied vacuously, because /t/ is –VOICED. Therefore, the output is [t]. In (8b), priority union adds the valued feature –VOICED to /D/, yielding [t]. In (8c), priority union yields [d] because /d/, which is +VOICED, is the strict argument, and –VOICED is a member of the defeasible argument. Priority union thus allows us to make a rule that changes underlying /D/ without affecting /d/ or /t/.

In order to use priority union to account for feature-changing processes, we need to dissociate two aspects of rule notation. In a traditional rule like ‘ $a \rightarrow b / \_ c$ ’ the first element  $a$  is called the *target*, and the rule is understood to apply to strings that contain the subsequence  $ac$ , turning that subsequence into  $bc$ . The element  $b$  is called the *structural change* of the rule. We need to maintain this convention for expressing rules, but distinguish it from the issue of which argument of priority union is the strict one and which is the defeasible one. This information is expressed by the direction of the arrow in the priority union symbol. Consider the examples of rules using priority union in (9).

(9) Rule target vs. strict argument

- i.  $a \cup \rightarrow b / \_ c$   
 $a$  is the target;  $a$  is the strict argument
- ii.  $a \cup \leftarrow b / \_ c$   
 $a$  is the target;  $b$  is the strict argument

In (9i),  $a$  is both the target of the rule and the strict argument of priority union. In (9ii),  $a$  is again the target of the rule, but now  $b$  is the strict argument of priority union. In both rules, the environment is defined by  $a$  being followed by  $c$ .

An illustration of feature changing rules does not require an underlyingly underspecified segment such as /D/, therefore a language with just /t/ and /d/ such as Hungarian is appropriate. Hungarian has a process of reciprocal neutralization whereby /t/ becomes

[d] before a voiced obstruent and /d/ becomes [t] before a voiceless obstruent. The process generalizes to all pairs of voiced and voiceless obstruents, such as /p/ and /b/ (with some irrelevant complications). For example, the forms in (10) show stem-final /p, t/ mapping to [b, d], respectively, before a suffix that begins with voiced /b/; and /b, d/ mapping to [p, t], respectively, before a suffix that begins with voiceless /t/.

(10) Hungarian feature-changing voicing assimilation

Lexical	Bare Noun	in Noun /-ban/	from Noun /-to:l/	to Noun /-nak/	gloss
/ku:t/	ku:t	ku:dban	ku:ttol	ku:tnak	‘well’
/ka:d/	ka:d	ka:dban	ka:ttol	ka:dnak	‘tub’
/kalap/	kalap	kalabban	kalapto:l	kalapnak	‘hat’
/rab/	rab	rabban	rapto:l	rabnak	‘prisoner’

This process can be captured with a single feature-changing rule that uses priority union and  $\alpha$ -notation, as shown in (11).

(11) Hungarian voicing assimilation

Examples:

- i.  $p, t \rightarrow b, d / \_\_ b$
- ii.  $b, d \rightarrow p, t / \_\_ t$

Rule:  $[-\text{SON}] \cup \leftarrow \{\alpha\text{VOICED}\} / \_\_ [\alpha\text{VOICED}, -\text{SON}]$

The target of the rule in (11) is the natural class of segments that are specified  $-\text{SON}$ , or more formally, the set of segments that are supersets of the set  $\{-\text{SON}\}$ , so this specification is listed leftmost in the rule, within square brackets. Next comes the priority union symbol, but with the arrow pointing from the right to the left, showing that the target’s

values are defeasible. Notice that the feature of the structural change ‘comes from’ (via  $\alpha$ -notation) the segment that defines the triggering environment. This is what it means to be an assimilation rule. Of course, the structural change can also be stipulated, without an alpha variable. A rule that devoices obstruents word-finally, as in Catalan, stipulates that the structural change involves priority union with the strict argument  $\{-VOICED\}$ , but this value is not bound to any segment in the triggering environment. In fact, no segmental trigger is necessary—a language could have feature-changing devoicing in word-final position, as exemplified by (12).

(12) Feature-changing final devoicing

$$[-SON] \cup_{\leftarrow} \{-VOICED\} / \_\%$$

Compare the rule in (12) to the feature-filling devoicing in (6) where the arrow of priority union points from the target towards the structural change. The rule in (12) is feature-changing because the voicing values on the targets are defeasible.

This example demonstrates the concern of Logical Phonology with ‘deconstructing’ the arrow of traditional phonological rules to better understand what kinds of functions phonological competence can contain. In other work (e.g., Bale et al. 2014, Volenec 2019 etc.), we have treated feature-filling using unification, which is a partial operation that, like priority union, is also related to set union. There, following a long tradition dating at least to Harris (1984) and adopted by many other scholars (such as Poser, 1993, 2004; Wiese 2000; Samuels 2011; McCarthy 2007) feature changing rules are modeled as a two-step process of deletion, accomplished with set subtraction, followed by insertion (i.e., feature-filling), accomplished with unification. The exploration of an alternative based on priority union reflects the methods and goals of Logical Phonology to understand the formal infrastructure of the language faculty.



#### 4 THE PHONOLOGY-PHONETICS INTERFACE: COGNITIVE PHONETICS

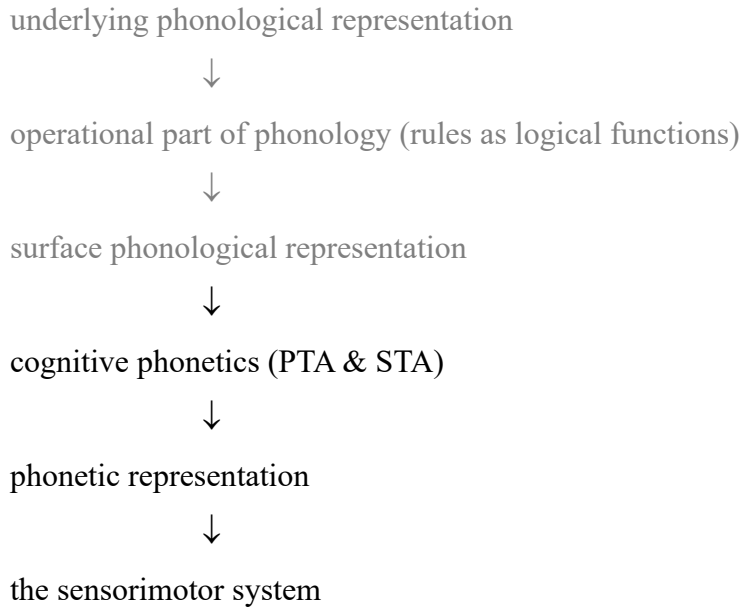
From the preceding discussion it is apparent that phonology is vastly different from phonetics. Phonology consists of abstract, symbolic, discrete, timeless units and formal operations, while phonetics is characterized by gradient, temporally arranged articulatory movements and continuously varying sound waves. Since humans do speak, it is a logical necessity that phonological competence successfully relays information to the sensorimotor (SM) system<sup>17</sup> which is in charge of speech production. The communication between these two systems is the *phonology-phonetics interface* (PPI).

*Cognitive Phonetics* (CP; Volenec & Reiss 2017; 2019) proposes that the PPI consists of two transduction procedures that convert the substance-free output of phonology into a representational format that contains substantive information required by the SM system in order 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 (still abstracting away from prosodic units). Each feature of SRs is transduced and subsequently receives interpretation by the SM system (cf. Lenneberg 1967: §3). This transduction is carried out by two algorithms (cf. Marr 1982/2010: 23–24). The *paradigmatic transduction algorithm* (PTA) takes a feature (a symbol in the brain) and relates it 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),<sup>18</sup> and it can be defined as a complex array of temporally

<sup>17</sup> Since the term *sensorimotor system* is used ambiguously in the linguistic literature, as was mentioned in the introductory section, it should be noted that we use that term to denote a large-scale brain network that includes the pre-central and post-central gyrus and the supplementary motor area.

<sup>18</sup> In generative linguistics literature, the output of phonological computation, what we call a *surface*

coordinated neuromuscular commands that activate muscles involved in speech production. The standard schema of phonological competence presented in section 2 can now be expanded to accommodate the transduction performed by CP:



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.

To clarify the effects of PTA and STA, we can explore in some detail the transduction of a few hypothetical SRs. Interestingly, as will be shown, 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 transduced phonological

---

*phonological representation*, is sometimes called a *phonetic representation*. Note that we use the term very differently—our phonetic representations do not contain the phonological features that are in URs and SRs.

features. Suppose that a hypothetical I-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. 2009: 719–720)—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,  $PR_{[+ROUND]}$  is the transduced feature +ROUND. We can now say that  $PR_{[+ROUND]}$  will be different for [o] because of its interaction with  $PR_{[-HIGH]}$  than for [u] because of its interaction with  $PR_{[+HIGH]}$ . Since these interactions involve transduced features within a single segment, [o] or [u], we can refer to these effects as *intrasegmental 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  $PR_{[+ROUND]}$  from the vowel onto the preceding consonant, i.e., in the anticipatory direction. This amounts to the more familiar *intersegmental coarticulation*, where transduced features from *different* segments interact. Returning to SRs [lok] and [luk], it is now apparent that (a)  $PR_{[+ROUND]}$  is different for [o] than for [u] due to its intrasegmental coarticulation with  $PR_{[HIGH]}$ ; (b) [l]’s inherent  $PR_{[-ROUND]}$  is now temporally overlapping with the  $PR_{[+ROUND]}$  from the adjacent vowels because of intersegmental coarticulation. It is important to note that the difference in  $PR_{[+ROUND]}$  from [o] and  $PR_{[+ROUND]}$  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]. Thus, [l] simultaneously bears the effect of both intra- and 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 *deterministic*, which means that it assigns the same neuromuscular schema 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.

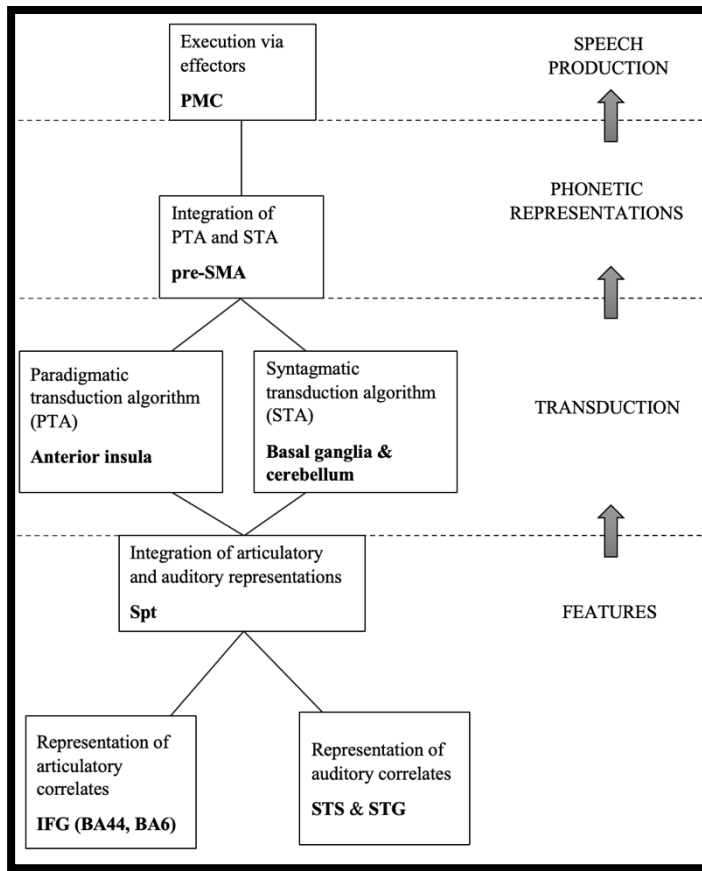
It should be stressed that CP's outputs, phonetic representations, should not be equated with actual articulatory movements or with the acoustic output of the human body. What is actually pronounced is further complicated in the process of externalization by a great number of factors. Transduction is followed by other performance factors that have no bearing on either phonology or transduction, factors like loudness, muscle fatigue, degree of enunciation, interruptions due to sneezing, 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 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. The apparent *lack of invariance* in the realization of a cognitively invariant category is not a matter of transduction. In fact, lack of invariance is only problematic for those scholars of language and speech who assume a misguided *physicist* stance (as opposed to a mentalist stance) towards their respective objects of study. Philosopher Irene Appelbaum (2006) has pointed out that “nearly half a century of sustained effort in a variety of theoretical perspectives has failed to solve the problem” of the lack of invariance in speech perception—in the output of the human body there are no constant correlates of particular segments or features across contexts. Appelbaum

suggests that scholars attempting to find such invariant properties may be asking the wrong questions, and we find, in the study of vision for example, that the assumption that perception is stimulus bound must be rejected (Marr 1982/2010). This was recognized in the study of language and speech already by Sapir (1933/1949: 46): “No entity in human experience can be adequately defined as the mechanical sum or product of its physical properties”. In contrast, assuming a mentalist stance towards the study of language and speech, and appreciating the fact that it is phonetics that is grounded in phonology and not the other way around, leads us to perceive the lack of invariance not as a ‘problem’ but rather as an expected prediction of our theory (i.e., it follows naturally from the competence/performance dichotomy). Performance will always introduce into the acoustic record some degree of completely random and chaotic component, thus even in acoustic phonetics a certain degree of abstraction is necessary to arrive at law-like conclusions.

Recent neuroscience evidence is consistent with the idea that CP transduces features, which are to be understood as symbols in the brain, into temporally distributed neuromuscular activities (elements of PRs), relating phonological competence to the vastly different SM system (Dronkers 1999; Hickok & Poeppel 2007; Guenther et al. 2006; Eickoff et al. 2009, Hickok 2012).<sup>19</sup> The activity in parts of the inferior frontal gyrus (IFG) corresponds to the representations of the articulatory correlates of features, while the activity in parts of the superior temporal gyrus (STG) and sulcus (STS) corresponds to the representations of the auditory correlates of features. An area in the Sylvian fissure at the boundary between the parietal and the temporal lobe (Spt) unifies these two

---

<sup>19</sup> See Volenec & Reiss (2017: §4.3) for a more elaborate exposition of the neural basis of Cognitive Phonetics, which also includes extensive referencing to previous neuroscience work that corroborates it.



**FIGURE 2.** NEURAL SUBSTRATE OF PHONOLOGICAL FEATURES AND OF THE PHONOLOGY-PHONETICS INTERFACE.

aspects into a complete symbol, a feature. The symbols are sent to the anterior insula where the PTA is carried out, and to the cerebellum and the basal ganglia where the STA is carried out. The PTA and the STA are integrated in the anterior part of the supplementary motor area (pre-SMA) 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. These neural processes are graphically represented in Figure 2.

Since they are part of performance and not of competence, the transduction algorithms cannot be specific to any particular I-language, that is, they are *universal*. However, although they are biologically universal in humans, CP will still show ‘output variability’ due to the PTA and the STA being applied to SRs that are featurally distinct. It is therefore important to distinguish between the outputs of the transduction system (phonetic representations) and the system itself (CP). The outputs of the system are, trivially, I-language-dependent and surface-representation-dependent since CP will automatically transduce any SR that is fed to it and any particular I-language will generate its own distinct set of SRs. For example, English URs and SRs will not have click segments in them, but some !Xhosa URs and SRs

will have clicks. For CP, it is irrelevant whether an SR comes from an English, !Xhosa or Telugu speaker. The CP system itself is part of the human biological make-up and can in this sense be regarded as a universal property of the human species. The universality of CP is merely a reflection of the fact that there exists a biological object we may informally refer to as *the phonetic implementational system*, of which CP is one part and the SM system another.

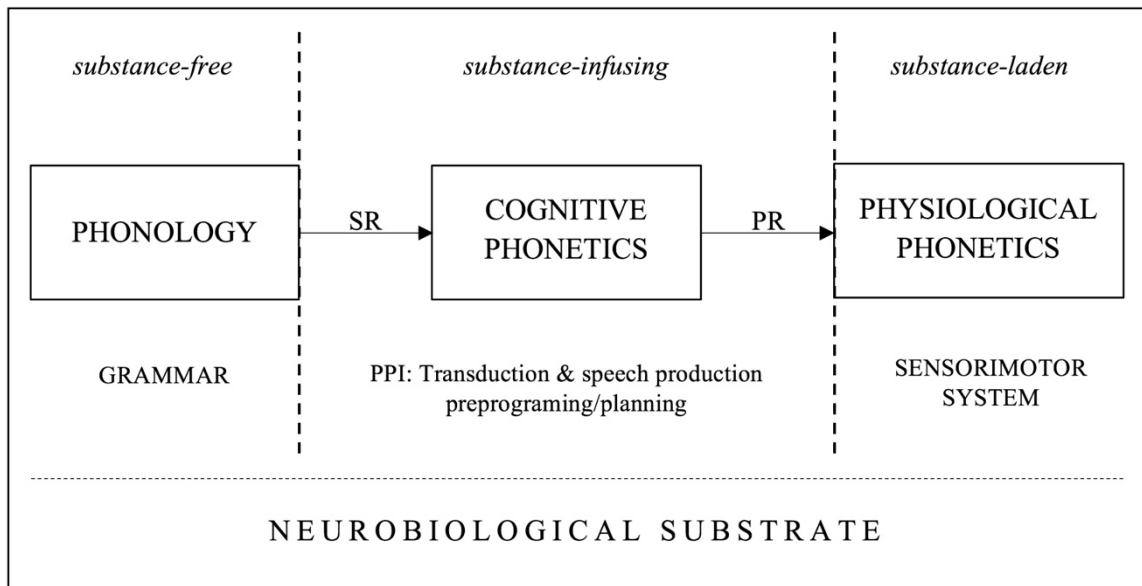
Consequently, we deny the existence of ‘language-specific phonetics’. This is simply a result of consistently adhering to the competence/performance dichotomy. Something can be language-specific only by virtue of belonging to an individual’s linguistic competence, i.e., to their I-language. Something that is not part of an I-language cannot be specific to an I-language, i.e., it cannot be language-specific. Since speech, the object of study of phonetics, is not language (competence, I-language), no part of it can be language-specific. The idea of ‘language-specific phonetics’ is thus a contradiction in terms just as, say, the idea of ‘language-specific vision’, because neither vision nor speech are part of linguistic competence.<sup>20</sup> Adherence to *universal phonetics* means that all alleged ‘language-specific phonetic patterns’ have to be reanalyzed by capturing them through phonological features and other phonological primitives,<sup>21</sup> while at the same time strictly resisting the temptation of lumping distinct I-languages under socio-politically determined labels (e.g. ‘English language’, ‘General American’, ‘German dialects’). That this feat is indeed possible is reflected in the fact that 30 binary features (slightly more than is usually assumed) with underspecification can yield over 206 trillion distinct segments, which is more than enough to capture all purported examples of ‘language-specific phonetics’, such as the systematic difference in the high front unrounded vowel between

<sup>20</sup> Note that both the visual system and the articulatory-auditory system *interface* and *communicate* with linguistic competence—vision in reading, the articulatory-auditory system in speaking—but neither is *part of* competence (see Volenec & Reiss 2017 for details).

<sup>21</sup> Of course, the features that are currently assumed by most phonologists, presented with some variation in modern textbooks such as Hayes (2009), Odden (2013), Gussenhoven & Jacobs (2017) etc., might not be enough to achieve this task and a slight enrichment of the universal feature set might be necessary. See Hale et al. (2007) for an elaboration along this line.

English and German speakers. In our framework, a German speaker's [i] and an English speaker's [i] must actually correspond to different feature sets because there is no language-specific phonetics which could have introduced a difference. Given the combinatoric explosion discussed in section 2, it is to be expected that there are many different feature sets (segments) that inhabit the 'high front unrounded vowel' space.

Figure 3 summarizes the general architecture of the phonology-phonetics interface (PPI) in the theory of Cognitive Phonetics (CP). To connect substance-free phonology with the substance-laden physiological phonetics, CP takes features of phonological surface representations (SRs) and relates them to neuromuscular activity (PTA) and arranges that activity temporally (STA), thus generating phonetic representations (PRs) that are directly interpretable by the sensorimotor (SM) system.



**FIGURE 3.** THE ARCHITECTURE OF THE PHONOLOGY-PHONETICS INTERFACE AND THE PLACE OF COGNITIVE PHONETICS WITHIN IT.



## CONCLUSION

The object of study in linguistics is I-language, a module of an individual's mind/brain that generates an unbounded array of discrete hierarchical expressions. I-language is further divided into distinct submodules, one of which is phonology (phonological competence). Phonological representations are built from certain atomic symbols, such as the features (VOICED, ROUND, etc.) and values (+, –); and phonological computation involves rules built from logical functions such as priority union. The entirety of phonology resides in the human mind/brain, and its connection to phonetic substance (e.g., movements of the articulators during speech) is remote and complex. Given its mind-internal nature, phonology is to be studied consistently as a branch of cognitive science, seeking ultimate unification with neurobiology, with all the corresponding theoretical and methodological commitments that follow from this mentalistic perspective.

In this paper, we have provided a general program for the study of phonology as a branch of cognitive science. Resting on the philosophical foundations of generative linguistics (Chomsky 1968; 1986; 2000), this program can be called *Formal Generative Phonology*. Its main tenet is that phonological competence is to be characterized as a formal—that is, explicit, logically precise, and substance-free—manipulation of abstract symbols. We have suggested that a productive way to execute this program is to adopt a model of formal generative phonology called *Logical Phonology* (building on work like Bale & Reiss 2018). There, phonological competence is described and explained by the application of set-theoretic functions, such as subtraction, union, generalized intersection, priority union etc., to abstract representational data structures such as strings of segments. The remote and complex relationship between phonological competence and speech was then elucidated by *Cognitive Phonetics* (Volenec & Reiss 2017; 2019), which proposes that the outputs of phonology are transduced via two algorithms into temporally distributed neuromuscular activities.

A foundational pillar of linguistics as a modern science is the idea that language

and speech are two different but connected entities. Over one hundred years ago, Ferdinand de Saussure likened their relationship to two different sides of a single sheet of paper, arguing persuasively that the formal underlying system (*langue*) must be strictly distinguished, both in theory-construction and in analytical practice, from the substance-based use of that system in communication (*parole*):<sup>22</sup>

*Language is a form and not a substance.* This truth could not be overstressed, for all the mistakes in our terminology, all our incorrect ways of naming things that pertain to language, stem from the involuntary supposition that the linguistic phenomenon must have substance. (Saussure 1916/1959: 122)

A century later, the most serious mistakes of phonology still stem from the supposition that phonological phenomena must have phonetic substance. If one thinks that phonology is language and not speech, then one is compelled also to accept that phonology is form and not substance. This line of reasoning provides “a principled manner” for determining “the boundary between transducer and symbolic processor” (Pylyshyn 1984: 148), or in more familiar terms, between phonetics and phonology. And with this conceptual boundary in place, we are in a better position to further develop an *explicit* theory of phonological representation and computation. We have provided a general program for such a phonological enterprise, a program that is grounded in cognitive neuroscience and formal logic, and not in phonetic substance and intuition-based claims. Is such a program on the right track? “Time will tell. But intuition won’t.” (Bromberger & Halle 1995: 739)

<sup>22</sup> A well-known difference between Saussurean and generative linguistics is that in the former language is not a mental object while in the latter it is. The important point here is that De Saussure’s *form* vs. *substance* dichotomy is valid in a completely mentalistic, internalist approach to language.

## REFERENCES

- Akiyama, K. et al. (2019). First M87 Event Horizon Telescope Results. I. The Shadow of the Supermassive Black Hole. *The Astrophysical Journal Letters*, 875/1: 1–17.
- Appelbaum, I. (1996). The lack of invariance problem and the goal of speech perception. *Proceeding of Fourth International Conference on Spoken Language Processing*, Vol. 3.
- Archangeli, D. & D. Pulleyblank (2015). Phonology without universal grammar. *Frontiers in Psychology*, 6.
- Bale, A., M. Papillon & C. Reiss. (2014). Targeting underspecified segments: A formal analysis of feature changing and feature filling rules. *Lingua*, 148: 240–253.
- Bastiaanse, R. & Thompson, C. (2003). Verb and auxiliary movement in agrammatic Broca's aphasia. *Brain and Language*, 87: 286–305.
- Bastiaanse, R. & van Zonneveld, R. (1998). On the relation between verb inflection and verb position in Dutch agrammatic aphasics. *Brain and Language*, 64: 165–81.
- Bergelson, E. & Swingle, D. (2012). At 6–9 months, human infants know the meanings of many common nouns. *Proceedings of the National Academy of Sciences*, 109(9), 3253–3258.
- Binder, J. R., Desai, R. H., Graves, W. W., & Conant, L. L. (2009). Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cerebral Cortex*, 19(12), 2767–2796.
- Blank, I., Kanwisher, N., & Fedorenko, E. (2014). A functional dissociation between language and multiple-demand systems revealed in patterns of BOLD signal fluctuations. *Journal of neurophysiology* 112/5: 1105–1118.
- Blevins, J. (2004). *Evolutionary phonology: The emergence of sound patterns*. Cambridge: Cambridge University Press.
- Bloomfield, L. (1926). A Set of Postulates for the Study of Language. *Language*, 2/3: 153–164.
- Bottini, G., Corcoran, R., Sterzi, R., Paulesu, E., Schenone, P., Scarpa, P., & Frith, D. (1994). The role of the right hemisphere in the interpretation of figurative aspects of language: A positron emission tomography activation study. *Brain*, 117(6), 1241–1253.
- Bouchard, K. E., N. Mesgarani, K. Johnson & E. F. Chang. 2013. Functional organization of human sensorimotor cortex for speech articulation. *Nature*, 495/7441: 327–332.
- Brannon, E. M. (2005). The independence of language and mathematical reasoning. *Proceedings of the National Academy of Sciences*, 102(9), 3177–3178.
- Bromberger, S. & M. Halle (1995). The Ontology of Phonology. *Studies in Logic and the Foundations of Mathematics* 134: 725–743.
- Buchert, R., Thomasius, R., Wilke, E., Petersen, K., Nebeling, B., Obrocki, J., Schulze, O., & Schmidt, U. (2008). Sustained effects of ecstasy on the human brain: A prospective neuroimaging study in novel users. *Brain*, 131/11: 2936–45.
- Burton, M. W, Small, S., & Blumstein, S. E. (2000). The role of segmentation in phonological processing: An fMRI investigation. *Journal of Cognitive Neuroscience* 12: 679–690.
- Caramazza, A. (1988). Some aspects of language processing as revealed through the analysis of acquired

- aphasia: The lexical system. *Annual Review of Neuroscience*, 11: 395–421.
- Carr, P. (2006). Universal grammar and syntax/phonology parallelisms. *Lingua*, 116(5): 634–656.
- Champagne-Lavau, M. & Joannette, Y. (2009). Pragmatics, theory of mind and executive functions after a right-hemisphere lesion: Different patterns of deficits. *Journal of Neurolinguistics*, 22: 413–26.
- Charles, R. (2019). *Epistemic boundedness and poverty of the stimulus: Why phonology is not different*. Ms. Concordia University.
- Chomsky, N. (1951/1979). *Morphophonemics of Modern Hebrew*. New York: Routledge.
- Chomsky, N. (1955/1975). *The Logical Structure of Linguistic Theory*. New York: Plenum Press.
- Chomsky, N. (1957a). Review of “Fundamentals of Language” by Roman Jakobson and Morris Halle. *International Journal of American Linguistics*, 23/3: 234–242.
- Chomsky, N. (1957b) Review of “A Manual of Phonology” by Charles F. Hockett. *International Journal of American Linguistics*, 23/3: 223–234.
- Chomsky, N. (1962). Explanatory models in linguistics. In Nagel, E., P. Suppes & A. Tarski (eds.), *Logic, Methodology, and Philosophy of Science*. 528–550. Stanford, CA: Stanford University Press.
- Chomsky, N. (1964). *Current Issues in Linguistic Theory*. The Hague: Mouton.
- Chomsky, N. (1965). *Aspects of the Theory of Syntax*. Cambridge: MIT Press.
- Chomsky, N. (1968). *Language and Mind*. Cambridge: Cambridge University Press.
- Chomsky, N. (1976). On the biological basis of language capacities. In Rieber, R. W. (ed.), *The Neuropsychology of Language*. 1–24. Boston, MA: Springer.
- Chomsky, N. (1980). *Rules and representations*. New York: Columbia University Press.
- Chomsky, N. (1986). *Knowledge of Language. Its Nature, Origins, and Use*. New York: Praeger.
- Chomsky, N. (2000). *New Horizons in the Study of Language and Mind*. Cambridge: Cambridge University Press.
- Chomsky, N. (2007). Approaching UG from below. In Sauerland, U. & H. M. Gärtner (eds.), *Interfaces + recursion = language? Chomsky’s minimalism and the view from syntax-semantics*. 1–24. Berlin: Mouton de Gruyter.
- Chomsky, N., M. Halle & F. Lukoff (1956). On Accent and Juncture in English. In Halle, M. (ed.), *For Roman Jakobson: Essays on the Occasion of His Sixtieth Birthday*. 65–80. The Hague: Mouton.
- Chomsky, N. & M. Halle (1968). *The Sound Pattern of English*. New York: Harper & Row.
- Chomsky, N., A. J. Gallego & D. Ott (2017) Generative Grammar and the Faculty of Language: Insights, Questions, and Challenges. Ms., MIT, Universitat Autònoma de Barcelona, and University of Ottawa. [<https://ling.auf.net/lingbuzz/003507>].
- Corina, D., Poizner, H., Bellugi, U, Feinberg, T, Dowd, D., and O’Grady-Batch, L. (1992). Dissociations between linguistic and nonlinguistic gestural systems: A case for compositionality. *Brain and Language*, 43: 414–47.
- Curtiss, S. (1982). Developmental dissociations of language and cognition, in L. Obler and L. Menn (eds), *Exceptional Language and Linguistics*. New York: Academic Press, 285–312.
- Curtiss, S. (1988a). The special talent of grammar acquisition, in L. Obler and D. Fein (eds), *The*

*Exceptional Brain*. New York: The Guilford Press, 364-86.

- Curtiss, S. (1988b). Abnormal language acquisition and grammar: Evidence for the modularity of language, in L. Hyman and C. Li (eds), *Language, Speech, and Mind: Studies in Honor of Victoria A. Fromkin*. New York: Routledge, Kegan, & Paul, 81-102.
- Curtiss, S. (1995). Language as a cognitive system: Its independence and selective vulnerability, in C. Otero (ed.), *Noam Chomsky: Critical Assessments, 4*. New York: Routledge.
- Curtiss, S. (2011). Revisiting modularity, in Yukio Otsu (ed.), *Proceedings of the 11th Annual Tokyo Conference on Psycholinguistics*. Tokyo: Hituzi Syobo, 1-33.
- Dahl, Veronica. (2002). On implicit meanings. In *Computational logic: Logic programming and beyond*. 506–525. New York: Springer.
- Dehaene-Lambertz, G. & Pena, M. (2001). Electrophysiological evidence for automatic phonetic processing in neonates. *NeuroReport* 12/14: 3155–3158.
- Dehaene-Lambertz, G., Baillet, S. & Hertz-Pennier, L. (2002). Functional neuroimaging of speech perception in infants. *Science* 298: 2013–2015.
- Dronkers, N. F. (1996). A new brain region for coordinating speech articulation. *Nature*, 384/6605: 159–161.
- Eickhoff, S. B., S. Heim, K. Zilles & K. Amunts. (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.
- Emmorey, K. (2002). *Language, Cognition and the Brain: Insights from Sign Language Research*. Mahwah, NJ: Lawrence Erlbaum.
- Emmorey, K., Kosslyn, S. M., & Bellugi, U. (1993). Visual imagery and visual-spatial language: Enhanced imagery abilities in deaf and hearing ASL signers. *Cognition*, 46(2), 139-181.
- Fant, G. (1969). Distinctive features and phonetic dimensions. *STL-QPSR*, 10/2–3): 1–18.
- Fodor, J. A. (1980). Reply to Putnam. In M. Piattelli-Palmarini (ed.), *Language and Learning: The Debate Between Jean Piaget and Noam Chomsky*. 325–334. Cambridge, MA: Harvard University Press.
- Friederici, A. D. (2017). *Language in our brain: The origins of a uniquely human capacity*. Cambridge: MIT Press.
- Gallistel, R. C. & A. P. King (2010). *Memory and the Computational Brain. Why Cognitive Science Will Transform Neuroscience*. Malden: Wiley-Blackwell.
- Gernsbacher, M. A., Morson, E. M., & Grace, E. J. (2016). Language and speech in autism. *Annual review of linguistics*, 2, 413-425.
- Greenberg, J. (1966). *Universals of language*. Cambridge, MA: MIT Press.
- Grinstead, J., MacSwan, J., Curtiss, S., & Gelman, R. (1997). The independence of language and number. In *Twenty-Second Boston University Conference on Language Development*.
- Grodzinsky, Y. (1986). Language deficits and the theory of syntax. *Brain and Language*, 27: 135–59.
- Grodzinsky, Y. & Finkel, L. (1998). The neurology of empty categories. *Journal of Cognitive Neuroscience*, 10/2: 281–92.
- Grodzinsky, Y. & Novogrodsky, R. (2006). Syntactic movement in agrammatism and S-SLI: Two different

- impairments, in A. Belletti, E. Bennati, C. Chesi, E. Di Domenico, & I. Ferrari (eds), *Language Acquisition and Development*. Newcastle: Cambridge Scholars Press, 197–210.
- Grosvald, M., Gutierrez, E., Hafer, S., & Corina, D. (2012). Dissociating linguistic and non-linguistic gesture processing: Electrophysiological evidence from American Sign Language. *Brain and language*, 121(1), 12–24.
- Guenther, F. H., Satrajit S. G. & J. A. Tourville. (2006). Neural modeling and imaging of the cortical interactions underlying syllable production. *Brain and Language*, 96/3: 280–301.
- Gussenhoven, C. & H. Jacobs (2017). *Understanding Phonology*. Fourth edition. London: Hodder Arnold – Hachette Livre.
- Hale, M. & C. Reiss. (2003). The Subset Principle in phonology: why the tabula can't be rasa. *Journal of Linguistics*, 39/2: 219–244.
- Hale, M., M. Kissock & C. Reiss. (2007) Microvariation, variation, and the features of universal grammar. *Lingua*, 117: 645–665.
- Hale, M. & C. Reiss. (2008). *The Phonological Enterprise*. Oxford: Oxford University Press.
- Halle, M. (1954). The strategy of phonemics. *Word* 10/2–3: 197–209.
- Halle, M. (1959). *The Sound Pattern of Russian*. The Hague: Mouton.
- Halle, M. (1961). On the Role of Simplicity in Linguistic Descriptions. *Proceedings of Symposia in Applied Mathematics* 12: 89–94.
- Halle, M. (1962). Phonology in Generative Grammar. *Word*, 18: 54–72.
- Halle, M. (1964). On the Bases of Phonology. In Fodor, J. A. & J. Katz (eds.), *The Structure of Language*. 324–333. New York: Prentice Hall.
- Halle, M. (1975). Confessio grammatici. *Language*, 51/3: 525–535.
- Halle, M. (1978/2002). Knowledge Unlearned and Untaught. In Halle (2002): 95–104.
- Halle, M. (2002). *From Memory to Speech and Back*. Berlin – New York: Mouton de Gruyter.
- Halle, M. & A. Marantz (1993). Distributed morphology and the pieces of inflection. In Hale, K. & S. J. Keyser (eds.), *The View from Building 20*. 111–176. Cambridge, MA: MIT Press.
- Hammarberg, R. (1976). The metaphysics of coarticulation. *Journal of Phonetics*, 4: 353–363.
- Harley, H & R. Noyer (1999). Distributed Morphology. *GLOT International*, 4/4: 3–9.
- Harris, J. W. (1984). Autosegmental phonology, lexical phonology and Spanish nasals. In M. Aronoff & R. Oehrle (eds.), *Language sound structure*. 67–82. Cambridge, MA: MIT Press.
- Hart, J., Berndt, R. S., & Caramazza, A. (1985). Category-specific naming deficit following cerebral infarction. *Nature*, 316: 439–40.
- Hayes, B. (2009). *Introductory Phonology*. Oxford: Wiley-Blackwell.
- Heisenberg, W. (1962). *Physics and Philosophy: The Revolution in Modern Science*. New York: Harper & Row.
- Hickok, G. (2012). Computational neuroanatomy of speech production. *Nature Reviews Neuroscience*, 13/2: 135–145.

- Hickok, G. & D. Poeppel. (2007). The cortical organization of speech processing. *Nature Reviews Neuroscience*, 8/5: 393–402.
- Hillis, A. E. (2010). The ‘standard’ for poststroke aphasia recovery. *Stroke*, 1316–1317.
- Hirsh-Pasek, K., & Golinkoff, R. (1996). *The Origins of Grammar: Evidence from Early Language Comprehension*. Cambridge, MA: MIT Press.
- Hirsh-Pasek, K., & Golinkoff, R. M. (1996). The intermodal preferential looking paradigm: A window onto emerging language comprehension. In D. McDaniel, C. McKee, & H. S. Cairns (eds.), *Language, speech, and communication. Methods for assessing children's syntax*. 105–124. Cambridge, MA: The MIT Press.
- Householder, F. (1965). On Some Recent Claims in Phonological Theory. *Journal of Linguistics*, 1: 13–34.
- Inkelas, S. (1995). The consequences of optimization for underspecification. In *Proceedings of the North East Linguistic Society*, 25: 287–302. University of Pennsylvania: Graduate Linguistic Student Association.
- Idsardi, W. & E. Raimy (2013). Three Types of Linearization and the Temporal Aspects of Speech. In Roberts, I. & M. T. Biberauer (eds.), *Challenges to Linearization*. 31–56. Berlin: De Gruyter.
- Jackendoff, R. (1990). *Semantic Structures*. Cambridge, MA: MIT Press.
- Jackendoff, R. (2015). In defense of theory. *Cognitive science*, 41: 185–212.
- Jacob, P. (2010). Intentionality. In E. N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*. Available at: <http://plato.stanford.edu/entries/intentionality>.
- Jakobson, R., G. Fant & M. Halle. (1952). *Preliminaries to speech analysis*. Cambridge: MIT Press.
- Jodzio, K., Biechowska, D., & Leszniewska-Jodzio, B. (2008). Selectivity of lexical-semantic disorders in Polish-speaking patients with aphasia: Evidence from single-word comprehension. *Archives of Clinical Neuropsychology*, 23/5: 543–51.
- Kager, R. (1999). *Optimality Theory*. Cambridge: Cambridge University Press.
- Kaplan, R. M. (1987). Three seductions of computational psycholinguistics. In P. Whitelock, M. M. Wood, H. L. Somers, R. Johnson, & P. Bennett (eds.) *Linguistic theory and computer applications*. 149–188. London: Academic Press.
- Karmiloff-Smith, A., Klima, E., Bellugi, U., Grant, J., & Baron-Cohen, S. (1995). Is there a social module? Language, face processing, and theory of mind in individuals with Williams syndrome. *Journal of cognitive Neuroscience*, 7(2), 196–208.
- Keating, P. A. (1985). Universal phonetics and the organization of grammars. In Fromkin, V. (ed.), *Phonetic linguistics: essays in honor of Peter Ladefoged*. 115–132. Orlando: Academic Press.
- Keating, P. A. (1990). Phonetic representations in a generative grammar. *Journal of Phonetics*, 18: 321–334.
- Lenneberg, E. 1967. *Biological Foundations of Language*. New York: Wiley.
- Leonard, L. B. (2017). Phonological impairment. *The handbook of child language*, 573–602.
- Lust, B. (2006). *Child Language – Acquisition and Growth*. Cambridge: Cambridge University Press.
- MacGregor, L. J., Pulvermüller, F., Van Casteren, M., & Shtyrov, Y. (2012). Ultra-rapid access to words in the brain. *Nature Communications* 3: 711.

- 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, MA: MIT Press.
- Matamoros, C. & C. Reiss (2016). Symbol taxonomy in biophonology. In A. M. Di Sciullo (ed.), *Biological Investigations on the Language Faculty*. 41–54. Amsterdam: John Benjamins.
- McCarthy, J. J. (2002). *A Thematic Guide to Optimality Theory*. Cambridge: Cambridge University Press.
- McCarthy, J. J. (2007). Slouching toward optimality: Coda reduction in OT-CC. *Phonological Studies (Journal of the Phonological Society of Japan)*, 7: 89–104.
- McCarthy, J. J. & A. Prince (1995). *Faithfulness and Reduplicative Identity*. Rutgers Optimality Archive: <http://roa.rutgers.edu/files/216-0997/roa-216-mccarthy-1.pdf>.
- McGuire, P. K., Robertson, D., Thacker, A., David, A. S., Kitson, N., Frackowiak, R. S. J., & Frith, C. D. (1997). Neural correlates of thinking in sign language. *NeuroReport*, 8/3: 695–8.
- Medawar, P. B. (1960). *The Future of Man: The BBC Reith Lectures 1959*. London: Methuen and Company.
- Mesgarani, N., Cheung, C., Johnson, K., & E. F. Chang. (2014). Phonetic feature encoding in human superior temporal gyrus. *Science* 343(6174), 1006–1010.
- Michotte, A. & L. Burke (1951). Une nouvelle enigme de la perception: le “donné amodal” dans l’expérience sensorielle. *Proceedings of the 13th International Congress of Psychology*, 179–180.
- Mielke, Jeff. (2008). *The emergence of distinctive features*. Oxford: Oxford University Press.
- Miller, G. A. (1998). Interview. In M. Gazzaniga, R. Ivry, & G. Mangun, (eds.), *Fundamentals of cognitive neuroscience*. New York: Norton.
- Mountcastle, V. B. (1998). Brain science at the century's ebb. *Daedalus*, 127/2: 1–36.
- Odden, D. (2013). *Introducing Phonology*. Second edition. Cambridge: Cambridge University Press.
- 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: 1036–55.
- Pinker, S. (1994). *The Language Instinct*. New York, William Morrow.
- 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.
- Poser, W. J. (1993). Are strict cycle effects derivable? In S. Hargus & E. M. Kaisse (eds.), *Studies in lexical phonology*, volume 4. San Diego: Academic Press.
- Poser, W. J. 2004. *On the status of Chumash sibilant harmony*. Ms. University of Pennsylvania.
- Prince, A. & P. Smolensky (1993/2004). *Optimality Theory: Constraint Interaction in Generative Grammar*. Oxford: Blackwell.
- Pulvermüller, F., Shtyrov, Y., Hasting, A. S., & Carlyon, R. P. (2008). Syntax as a reflex: Neurophysiological evidence for early automaticity of grammatical processing. *Brain and language* 104/3: 244–253.
- Pylyshyn, Z. (1984). *Computation and Cognition*. Cambridge: MIT Press.



- Ralph, M. A. L., Jefferies, E., Patterson, K., & Rogers, T. T. (2017). The neural and computational bases of semantic cognition. *Nature Reviews Neuroscience*, 18(1), 42.
- Reiss, C. (2012). Towards a bottom-up approach to phonological typology. In A. M. Di Sciullo (ed.), *Towards a Biolinguistic Understanding of Grammar*. 169–191. Amsterdam: John Benjamins Publishing Company.
- Rohlf, M. (2016). Immanuel Kant. In *Stanford Encyclopedia of Philosophy*. Available at: <https://plato.stanford.edu/entries/kant/>.
- Rondal, J. (1995). *Exceptional Language Development in Down Syndrome*. Cambridge: Cambridge University Press.
- Samuels, B. (2011). *Phonological architecture: a biolinguistic perspective*. Oxford: Oxford University Press.
- Sapir, E. (1933/1949). The Psychological Reality of Phonemes. In Mandelbaum, D. G. (ed.), *Selected Writings of Edward Sapir in Language, Culture and Personality*. 33–45. Los Angeles: University of California Press.
- Saussure, F. de (1916/1959) *Course in General Linguistics*. New York: Columbia University Press.
- Schaller, S. (1995). *A Man Without Words*. Berkeley: University of California Press.
- 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.
- Scharinger, M., P. J. Monahan & W. J. Idsardi. (2012). Asymmetries in the processing of vowel height. *Journal of Speech, Language, and Hearing Research*, 55/3: 903–918.
- Seikel, A. J., W. D. King & D. G. Drumright. (2009). *Anatomy & Physiology for Speech, Language, and Hearing*. 4th edition. New York: Delmar, Cengage Learning.
- Skeide, M. A., Brauer, J., & Friederici, A. D. (2014). Syntax gradually segregates from semantics in the developing brain. *NeuroImage*, 100, 106–111.
- Sklar, L. (2000). *Theory and Truth: Philosophical Critique within Foundational Science*. Oxford: Oxford University Press.
- Smith, N. (2003). Dissociation and modularity: reflections on language and mind. *Mind, Brain, and Language: Multidisciplinary Perspectives*, 87–111.
- Smith, N. (2010). *Acquiring phonology: a cross-generational case-study*. Cambridge: Cambridge University Press.
- Smith, N. V., & Tsimpli, I. M. (1995). *The mind of a savant: Language learning and modularity*. New York: Blackwell Publishing.
- Smith, N., Tsimpli, I., Morgan, G., & Woll, B. (2010). *The signs of a savant: Language against the odds*. Cambridge: Cambridge University Press.
- Tippett, D. C., & Hillis, A. E. (2016). Vascular Aphasia Syndromes. In *Neurobiology of Language* (pp. 913–922). Academic Press.
- Versace, E., Martinho-Truswell, A., Kacelnik, A., & Vallortigara, G. (2018). Priors in animal and artificial intelligence: where does learning begin? *Trends in cognitive sciences*, 22/11: 963–965.
- Volenec, V. (2019). Nasal place assimilation in Substance Free Logical Phonology. *Loquens* 3/1.

- Volenec, V. & C. Reiss (2017). Cognitive Phonetics: The Transduction of Distinctive Features at the Phonology-Phonetics Interface. *Biolinguistics*, 11: 251–294.
- Volenec, V. & C. Reiss (2019). The intervocalic palatal glide in Cognitive Phonetics. *Proceedings of the 49<sup>th</sup> Meeting of the North East Linguistics Society*. 255–264.
- Wang, P. P., Doherty, S., Rourke, S. B., & Bellugi, U. (1995). Unique profile of visuo-perceptual skills in a genetic syndrome. *Brain and Cognition*, 29(1), 54–65.
- Wiese, R. (2000). *The phonology of German*. Oxford: Oxford University Press.
- Willems, R. M., Özyürek, A., & Hagoort, P. (2009). Differential roles for left inferior frontal and superior temporal cortex in multimodal integration of action and language. *Neuroimage*, 47(4), 1992–2004.
- Zatorre, R. J., Evans, A. C., Meyer, E., & Gjedde, A. (1992). Lateralization of phonetic and pitch processing in speech perception. *Science* 256: 846–849.