3.8 THE PHONETICS-PHONOLOGY INTERFACE

Silke Hamann, University Düsseldorf

1 Introduction

To define the interface of phonetics and phonology, we first have to establish the topics of the two linguistic disciplines that are involved. While phonology deals with abstract, discrete categories encoding language-specific differences in meaning, phonetics is concerned with gradient, continuous, physical representations encoding articulatory and auditory details (see e.g. Pierrehumbert 1990 and Cohn 1993 for detailed discussions). The main question for the phonetics-phonology interface is how these two types of representations can be connected: How do we map discrete, abstract categories onto a phonetic parametric space, and vice versa? This question is often restricted to one direction of processing, namely the production side, where it is known under the name of "phonetic implementation": How are phonological structures that only exist in our brain converted into something in the real world?

Until the mid 1970s, the view prevailed that every phonological category (i.e., every segment, feature, syllable, etc.) has an invariant characteristic in the acoustic signal or in the articulation. This invariant trait was considered to be language-independent. The mapping from abstract phonological material onto the discrete phonetic realization was thus believed to be universal and automatic. The search for these invariant, language-independent phonetic characteristics of phonological categories caused a massive increase in experimental methods, which eventually gave rise to the field of *Laboratory Phonology* (see chapters 2.2 and 4.1). However, finding invariant traits turned out to be more problematic than anticipated.

Early cross-linguistic studies on the acoustics of segments showed already that there are no universal correlates of phonological categories, see e.g. the work on voicing distinctions by Lisker & Abramson (1964, 1970, and later). Such findings had a huge impact on the conception of the phonetics-phonology interface (see the discussion in Fromkin 1975 and Keating 1988): Instead of an automatic process, phonetic implementation turned out to be something that has to be learned and that makes up part of our grammatical knowledge. Later studies revealed that even within a language there is no single invariant characteristic of phonological categories. In addition to inter- and intra-speaker variation, experimental work showed that the realizations of a category depend largely on context and prosodic position of the category, on the lexical frequency of the word in which it occurs, and on gender, regional background and social affiliation of the speaker (for literature on all these factors see the volumes on Laboratory Phonology starting with Kingston & Beckman 1990).

How can speakers/listeners deal with all this variation? And is this large amount of variation reconcilable with the notion of an abstract representation at all? Instead of an answer to the mapping between phonology and phonetics, the experimental studies provided us with more questions. This situation led Kingston (2006: 401) quite recently to exclaim that there is still no consensus on what constitutes the phonetics-phonology interface. A scholar new to the field might wonder how this is possible after such a long time of extensive studies. The answer proposed in this article is that the nature of the interface cannot be unearthed by experimental studies alone. It depends to a considerable part on the theoretical assumptions we make, and on the aim we have in mind with our phonological and phonetic description. This has been observed already by Ladefoged (1972) almost forty years ago, who summarized that there is not one but *many different* phonetics-phonology interfaces, depending on the aim of the studies and the assumptions following from it.

The present article is structured as follows. Section 2 shortly introduces three topics that can be considered essential for the interface, namely phonological features, the distinction between phonetic and phonological processes, and phonetic grounding. Section 3 deals with assumptions on phonological and phonetic representations, their possible connection, and how these assumptions shape the interface. A model with an explicit formalization of the interface and its predictions are introduced in section 4. Section 5 concludes.

2 Recurring topics of the phonetics-phonology interface

In this section, we look at three topics that are important for the phonetics-phonology interface and can be found recurrently in the literature on the interface. These are: phonological features (§2.1), possible distinctions between phonetic and phonological processes (§2.2), and the phonetic grounding of phonological processes (§2.3). They are relevant to the definition of the interface for the following reasons: While features are often considered to constitute the phonetics-phonology interface per se, the other two topics have a considerable impact on the distinction between phonetics and phonology and therefore on the connection between the two modules.

2.1 Phonological features

Phonological features are the smallest unit in phonology (next to moras with their comparatively restricted use). Features are defined as having direct correlates in articulation, acoustics or audition. Very often, they seem to be nothing more than a phonetic characteristic. The feature [labial], for instance, is shared by segments articulated with the lips. For this reason, features can be considered the interface between phonetics and phonology: They connect a phonological category (the segment) with a phonetic realization (its articulation, its percept, or its acoustic form). The feature set that is most commonly used in present-day phonology goes back to the one proposed by Chomsky & Halle (1968) and has almost only articulatory correlates. This set has been modified by Clements (1985) to mirror the dependencies between articulators (e.g., coronals can be articulated either with the tongue tip or blade), resulting in a more realistic model of concrete articulations associated with phonological categories. Alternatives to these articulatory features and thus a purely articulatory definition of the interface are proposed e.g. by Jakobson, Fant & Halle (1952), who introduced a feature set with acoustic, auditory and articulatory correlates, or the elements proposed by Harris & Lindsey (1995), which are unary features with acoustic definitions.[1]

Besides their function as phonetic interpretation of segments, features are used to group segments in the description of phonological processes: Segmental classes that undergo a process, are the result of a process, or form the context of a process are

[1] For a detailed discussion of features and alternatives to the SPE feature set, the interested reader is referred to chapter 3.1.

referred to by features. The class of sounds undergoing final devoicing in Dutch and German, for instance, are obstruents, defined by the feature [–sonorant].

Three problems arise for the phonetics-phonology interface with respect to the phonetic interpretability of features. First, experimental studies showed that phonological categories, including features, do not have one invariant phonetic trait (see the discussion on page 1). The direct phonetic interpretation of features is thus restricted and has to be supplemented by additional information on their context- and language-specific realization. If features constitute the interface, a listing of all additional language-specific correlates of features seems necessary for a complete description of the interface. Second, features can be defined in articulatory, acoustic and perceptual terms, and while articulatory specifications seem more straightforward in the production (the traditional "phonetic implementation"), a perceptual or acoustic specification seems to be more adequate for the perception process. If features constitute the interface, the question has to be answered whether the interface is restricted to production, perception or includes both (this point is taken up again in §3.4 below). Third, the phonetic traits of features cannot always be united with the function of phonological grouping. Sometimes a group of segments in a language forms a class in a phonological process but does not share a phonetic characteristic (see e.g. Mielke 2004 for examples of such "unnatural" classes). If features are assumed to come with their phonetic characteristics and thus provide the phonetics-phonology interface, then a different set of features not based in the phonetics seems necessary to account for unnatural phonological classes.[2]

In sum, referring to phonological features is not sufficient to fulfil the task of defining the phonetics-phonology interface. If features are assumed to constitute (at least part of) the interface, and it seems reasonable to connect phonetic realizations to the smallest phonological units, our assumptions on these phonetic realizations and on the use of features have to be made explicit. The questions raised here indicate how this could be done.

2.2 How can we distinguish between a phonetic and a phonological process?

The decision whether a certain process is phonological or phonetic is not always as easy as it seems. Word-final obstruent devoicing, for instance, has been traditionally considered a phonological process: Obstruents (the class of [–sonorant] segments) change their specification from [+voice] to [–voice] in word-final position, resulting in neutralization between voiced and voiceless obstruents in this position. More recent findings show that this process of devoicing is not complete. There are small acoustic differences between devoiced and underlyingly voiceless obstruents which can be detected by native listeners, see e.g. Port & O'Dell (1985) for German and Warner, Jongman, Sereno & Kemps (2004) for Dutch. But if the output of this process are small, traceable remnants of voicing, can it still be considered phonological? The answer to this question depends on our definition of phonology. If we consider phonology to

[2] This later point was already raised by Ladefoged (1972), who argued that different types of phonological descriptions require different types of features: A cross-linguistic comparison is best made with a restricted, unified set of categories (such as universal, phonetically-based features), whereas a complete description of a single language is best made with partly phonological features without phonetic correlates (to account for unnatural classes), and partly with features that have very detailed phonetic correlates (to account for small phonetic but non-distinct contrasts).

operate on abstract and discrete representations (such as the feature [±voice]), as we have defined it at the beginning of this chapter, then an incomplete neutralisation cannot be part of phonology but must be assigned to phonetics. The slightly disturbing question then arises how many of the processes that have been traditionally considered phonological (as neutralisations) and have been described in textbooks of phonology really are phonological, or are only lacking the phonetic evidence to show their gradual nature.

In an alternative conception of phonology that allows gradient phonological representations, processes with non-discrete outputs can be considered phonological. Flemming (2001), a proponent of such a non-discrete approach to phonology, mentions a further possible problem for the strict division between phonological and phonetic processes in the traditional view: The two types of processes often seem to be the same, and cause a reduplication of description in phonology and phonetics. Flemming mentions phonetic coarticulation and phonological assimilation processes as an example for such a reduplication. Both involve the change of one segment caused by an adjacent segment; in one case this change is gradual, in the other case discrete. Flemming argues that treating both as caused by the same type of restriction (an effort-avoidance constraint) in the phonological grammar simplifies the grammatical description.[3] This non-discrete approach to phonology and possible problems with it are discussed further in section 3.2 below.

We can observe that only explicit criteria of what the output of a phonological process is allowed to look like let us decide whether a process is phonological or phonetic. The way we define our phonological module thus determines what we consider to be part of phonetics. This in turn determines our conception of the phonetics-phonology interface, as we will see in section 3 below.

2.3 Phonetic grounding of phonological processes

A third topic that is mentioned recurrently in studies on the phonetics-phonology interface is phonetic grounding. The urge to seek phonetic motivations for phonological processes is far from new (see e.g. Passy 1891) but has not lost its fascination, see more recent studies on phonetic explanations e.g. in *Natural Phonology* (Stampe 1972, Donegan & Stampe 1979), in *Grounded Phonology* (Archangeli & Pulleyblank 1994), and in the work on sound change by Ohala (e.g., 1981, 1983, 1992 and following). The reason why researchers try to phonetically explain phonological observations is the fact that large numbers of unrelated languages often show similar phonological behaviour due to phonetics. We find for instance that many languages with retroflex segments avoid sequences of retroflexes and high front vocoids, which has been explained as the avoidance of sequences with extremely dissimilar tongue positions (Flemming 2003, Hamann 2003). The same articulatory motivation has been applied to the cross-linguistic avoidance of sequences with a rhotic and the glide /j/ (Hall & Hamann to appear). But there are also perceptual motivations, which are more often employed to account for the preference of whole segment inventories. The best-known example is the so-called dispersion effect in vowel inventories. Five-vowel inventories, for instance, mostly consist of [i e a o u], which has been explained by the fact that

[3] Cohn (2007) provides a more detailed discussion and counterarguments to this reduplication problem.

these five vowels are perceptually most distinct, i.e. are least often confused with each other (Ladefoged 1975: 235-238; Lindblom 1986).

Phonetic explanations like these have often been located at the phonetics-phonology interface. Scholars, however, differ on the question whether such explanations should be treated in linguistic theory at all. Let us look first at the approach taken in *Optimality Theory* (Prince & Smolensky 1993 [2004], McCarthy & Prince 1993, henceforth: OT; for a detailed account see chapter 3.5). Most OT approaches use phonological markedness constraints that incorporate the phonetic motivation directly into the phonological account. The avoidance of sequences with a rhotic and the glide /j/, for instance, can be captured in a constraint *rj, which is based on the articulatory complexity of this sequence. Such markedness constraints have the advantage that the phonetic explanation is transparent in the phonological description, and that the typological variation we find in the languages of the world can be accounted for with a difference in ranking between the relevant markedness and faithfulness constraints (what is known in the OT-literature as 'factorial typology'). However, such accounts duplicate phonetic information in phonology (see the criticism voiced by e.g., Blevins 2004; Hume 2004; Ohala 2005; Haspelmath 2006). Others therefore propose that the phonetic basis of phonological processes should not be part of the linguistic theory at all, but should be sought "in a technically 'extralinguistic' domain, that of historical linguistics" (Hale 2000: 241/242). Haspelmath (1999) and Blevins (2004) also argue for an account of the emergence of typologically similar patterns via diachronic adaptation, as given in Ohala's work (1981 et sequ.). What these latter studies lack is a formal account of how exactly such diachronic changes occur, i.e. how the phonetics influences the acquisition process of a child in such a way that she acquires a phonological grammar that differs from that of her parents.

What the two approaches seem to have in common is the assumption that phonetic grounding can only be formalized in the phonological module of the grammar. Their assumptions on the nature of phonology guide their solution: If phonology is purely about discrete units, then we cannot formalize the phonetic grounding of processes in phonology (e.g., Ohala, Blevins, Haspelmath). If our phonological theory allows phonetic principles in the shape of markedness constraints, on the other hand, then we can formalize phonetic grounding in phonology (traditional OT accounts). But why does phonetic grounding have to happen in the phonology? If we have an explicit definition of the phonetics-phonology interface, then the connection between the phonological module and the phonetic one can account for the phonetic motivation of phonological processes without duplicating phonetics in the phonology or leaving out the phonetic motivation at all (a proposal along these lines is described in section 4 below).

Again an exact definition of what belongs to the phonological module and what to phonetics seems essential to the approaches presented here. In the following section we look at the phonetics-phonology divide in more detail, namely at the representations that can be postulated in both modules and their possible connections. These connections, and the formalisation of them, constitute the phonetics-phonology interface.

3 An |a| is an /a/ is an [a] is an 🧍a🧍? The question of representations

The focus of this section is the question how the number and types of representations that we assume in our grammar model influence the definition of the phonetics-phonology interface. We can discern roughly three views on phonetic and phonological representations. According to the first view, phonetic representations are very detailed and there are no permanent phonological representations, see §3.1. The second view assumes that the two types of representations are fairly similar, as discussed in §3.2, and the third that they are incommensurable, see §3.3. Section 3.4 deals with two questions that arise if we assume that phonetic representations are different from phonological forms: How many phonetic forms can we discern, and how do they connect to the phonological representations.

3.1 No stored phonological representations

The most extreme view on phonological representations is found in *Exemplar Theory* (Johnson 1997, 2005; Pierrehumbert 2001; the following is a crude summary, for a complete overview the reader is referred to chapter 4.2). Exemplar Theory originated in cognitive psychology as an approach to concept learning. It has been applied to phonology due to the findings in Laboratory Phonology that the acoustic signal varies to such a large amount that it seems unreasonable to assume speakers have stored abstract representations and the means (rules or constraints) to derive them. In addition, it allows the connection between phonetic realisations and gender, social status and other non-linguistic information on the speaker. Exemplar theorists do not postulate a separate phonological component of the grammar but assume instead that phonological generalisations are formed on-line across the stored phonetic exemplars when necessary. This process of generalisation is described in the following quote from Beckman (2003):

> It is abstracted away from sensory experience only to the extent that two utterances with values that are close enough to be indistinguishable along some dimension of the phonetic space will be registered as two instances of the same thing along that dimension. (p. 79)

This minimal view of representations is difficult to reconcile with the notion of a phonetics-phonology interface, because if there are no phonological representations we cannot formalise a connection between them and the phonetic forms. Exemplar Theory is therefore not further elaborated in this chapter.

3.2 Phonological and phonetic representations are commensurable

What does it mean to assume that phonetic and phonological representations are fairly similar? Hyman (2001) summarizes this quite fittingly by writing that both "phonetics and phonology utilize a common alphabet" (p. 143) because, as he goes on to argue, they use the same descriptive terms such as 'voiceless stops' and 'nasalized vowels'. The view that phonology and phonetics use fairly similar representations can be found in a large number of phonological studies.

Chomsky and Halle (1968; henceforth: SPE), for instance, propose that phonological rules change binary phonological features stepwise into non-distinct, scalar specifications (what they term phonetic features). They assume that intermediate stages have "representations of a highly mixed sort" (p. 66). These assumptions imply that phonological and phonetic representations are in some way comparable. According to SPE, the phonetics-phonology interface is then the point where abstract representations are complemented with (or replaced by) gradient details. This view is quite prominent in later generative work (e.g., in the assumption that features provide the phonetic realisation of segments language-independently, see Kenstowicz 1994 for an overview).

The idea that phonetic and phonological representations are commensurable can also be found in approaches that make no distinction between phonetic and phonological surface representations, and incorporate phonetic details directly into phonology. Such models are summarized here as *minimal functional approaches*,[4] and the representations and their relation(s) look as given in Figure 1.

Underlying phonological representation
|
[Surface phonological/phonetic representation]

Figure 1: The two representations in a minimal functionalist approach.

Three arguments have been brought forward in support of this approach. First, there is allegedly no clear-cut distinction between phonetic and phonological representations. This point is illustrated with a quote from Flemming (2001: 8f.):

> Phonetics and phonology are not obviously distinguished by the nature of the representations involved, [...] most of the primitives of phonological representation remain phonetically based, in the sense that features and timing units are provided with broadly phonetic definitions. This has the peculiar consequence that sound is represented twice in grammar, once at a coarse level of detail in the phonology and then again at a finer grain in the phonetics.

This quotation exemplifies the idea familiar from SPE that phonetic and phonological representations are commensurable. But instead of 'duplicating' information on speech sounds in phonetics and phonology and focussing on the phonological part as done in SPE and later generative work, Flemming proposes that we should treat phonetic and phonological representations as the same, implying that both are subject to the same grammar component. Their only difference then lies in their role: phonological representations are those distinctive contrasts that are chosen language-specifically amongst all possible phonetic contrasts.

The second argument brought forward for an incorporation of phonetic details into phonological representations is that phonetic and phonological processes often seem to be the same, see §2.2. Third, the phonetic grounding of phonological processes should be visible in the phonology, recall §2.3. In OT accounts, markedness constraints can refer to phonological structure (e.g., *CODA) or to phonetic detail (e.g., the effort minimization constraint *RETROFLEX in Flemming 1995/2002, or the MINDIST constraint by Padgett 2001 requiring perceptual dispersion). Such approaches can be found in the work by e.g. Flemming (1995/2002, 1997, 2001, 2003, 2004), Padgett (2001, 2003b,

[4] 'Minimal' is used to distinguish these approaches from other functional theories that also incorporate phonetic motivations of phonological processes into their formalism but follow strict modularity. For an example, see §4.

2003a), Kirchner (1997, 1998/2001, 1999/2004) and Steriade (1995, 1997, 1999, 2001a, 2001b) (see Gordon 2007 for an overview and Hale 2000 for criticism).

In such a minimal-functionalist approach it is impossible to define the phonetics-phonology interface, since there is no clear-cut distinction between phonetic and phonological representations, and both are handled in one grammar component. In the following section, we see approaches that distinguish phonetic from phonological representations, where the phonetics-phonology interface is thus the mapping between these representations.

3.3 Phonetic and phonological representations are inherently different

As illustrated in §3.1 with Hyman's quote, the idea that phonetic and phonological representations are commensurable is largely based on the fact that both forms are represented with the same alphabet, namely the IPA notation. In phonology, more precisely when depicting phonological forms, the IPA symbols stand for abstract representations, that is, segments, which themselves can be treated as shortcuts for feature bundles. In phonetic forms, IPA symbols can stand for articulatory gestures (e.g., a raised tongue tip), acoustic information (e.g., a high second and third formant), or for perceptual cues (e.g., energy in the mid-region of the spectrum). Phonological accounts, however, often fail to give an explicit listing of what notations like /a/ and [a] stand for.[5] As a result, researchers often seem to forget that these forms cannot be the same: an abstract phonological form with features and syllable boundaries cannot at the same time be a concrete articulatory gesture or an acoustic form. We will look now at alternative approaches with disparate representations for phonetics and phonology, as has been argued for by e.g. Pierrehumbert (1990: 377ff.).

If a strict division between abstract phonological representations and detailed phonetic representations is made, then it follows that phonological processes (in the phonological module) can only be described as a mapping between phonological representations, i.e. of phonological onto phonological ones. The input to this mapping is the lexical form with segmental and featural representations. The output, the surface phonological representation, involves predictable phonological material such as syllable boundaries or feet. It further follows that, if one wants to include phonetic representations (such as articulatory gestures or other phonetic details), they have to be assumed *in addition* to the phonological ones. These minimally three representations are illustrated in Figure 2.

Underlying phonological representation | Surface phonological representation	PHONOLOGY
|	
Phonetic representation(s)	PHONETICS

Figure 2: Phonological and phonetic representations in a modular approach.

[5] A similar point is made by Hale (2000: 243): "It is common in phonological circles to use the symbols of the [...] IPA in a systematically ambiguous manner."

In the literature we can find several phonological approaches that restrict themselves to the description of the two phonological representations. Since the present article is on the phonetics-phonology interface, we are, however, interested in those phonological approaches that include phonetic representations and make explicit statements on the connection between these different types of representations. There are two approaches in the phonological literature that fulfil these criteria. In the first, the connection between the surface phonological representation and the phonetic representation(s) is assumed to be automatic, in the second language-specific.

The first view is adopted by Hale (2000) and Hale & Kissock (2007). In order to distinguish three representations, they do the following. The traditional notation with slashes, as in /a/, are employed for an underlying representation. Brackets, as in [a], are used for what they call 'phonetic representation', which is the output of the phonological grammar or the form "at the end of a derivation" (Hale 2000: 243). Additionally, they assume a 'phonetic realization', which they define as the acoustic impression or the articulatory realization of a segment, and transcribe with little human bodies, e.g. 🧍a🧍. The latter is also called 'bodily output'. The mapping between the phonetic representation and the bodily output is performed by what they call 'transducers'. Transducers convert one type of representation into another, namely a featural, phonological one into a phonetic one and vice versa. In this respect transduction differs from phonological computation, where only the same representations are involved. Hale & Kissock (2007) differentiate two types of transducers, an auditory and an articulatory one. The auditory transducer converts the percept into phonological features whereas the articulatory transducer converts phonological features into what they term 'a gestural score'. The exact working of these conducers, i.e. a formalisation, is not provided. The two transducers are innate and invariant, and "identical in all humans" (Hale & Kissock 2007: 84). The connections between the surface phonological representation and the phonetic representations are thus automatic. Together with Hale & Kissock's assumption that phonological features are drawn from a universal set, this results in the same phonetics-phonology interface for every language; the only task for researchers working within this framework on the interface is then to describe the automatic mapping between surface form (or more specific: phonological features) and the phonetic form(s).

The second view, where the connection between surface phonological structure and the phonetic form is learned and therefore language-specific, is proposed by Boersma in his *Functional Phonology* model (1998), and in the follow-up model of *Bidirectional Phonology and Phonetics* (Boersma 2007, 2009). As we saw already for Hale and Kissock above, the use of a third representation requires the introduction of an additional notational convention. Boersma solves this by using pipes for the underlying phonological form, as in |a|, slashes for the surface phonological representation, as in /a/, and brackets for the phonetic form(s), as in [a]. As in Hale & Kissock's approach, both directions of processing are described in this model, and for this purpose the phonetic form is further distinguished into an articulatory form and an auditory form. The latter is directly connected to the phonological surface form, and this connection constitutes the phonetics-phonology interface. It is arbitrary, which means that the learning child has to acquire the relevant perceptual cues for each segmental category. Boersma employs the same formalism for the connection between auditory and surface form as he uses for the connection between surface and underlying form (ranked OT constraints, see §4 for details). This implies that both phonology and phonetics are treated as part of the grammatical knowledge.

3.4 How phonological and phonetic representations can connect: Possible views of the interface

Both proposals discussed in the previous section argue for phonetics and phonology as two separate modules, and the mapping of a phonetic representation onto a phonological one as the locus of the interface. Both Hale & Kissock's and Boersma's proposals also include two directions of processing: a comprehension direction (the task of the listener) and a production direction (the task of the speaker). This contrasts with the majority of phonological studies, who focus on the production process (see the overview in Boersma & Hamann 2009a). Even most phonologists working on the perceptual basis of phonological processes, such as Hume and Johnson (2001), Flemming (2004, 2005), Steriade (1995, 2001a), and Wright (2001), assume or explicitly claim that speech perception informs phonology but lies outside the scope of phonological theory (exceptions are Broselow 2004 and Pater 2004). This view of phonology restricts the phonetics-phonology interface to the connection between surface and articulatory representation, as illustrated in Figure 3. Note that the differentiation between surface and articulatory form in Figure 3 does not hold for all of the afore-mentioned studies, since most of them belong to the minimal functionalist approach elaborated in §3.1, and therefore collapse surface phonological and phonetic representations.

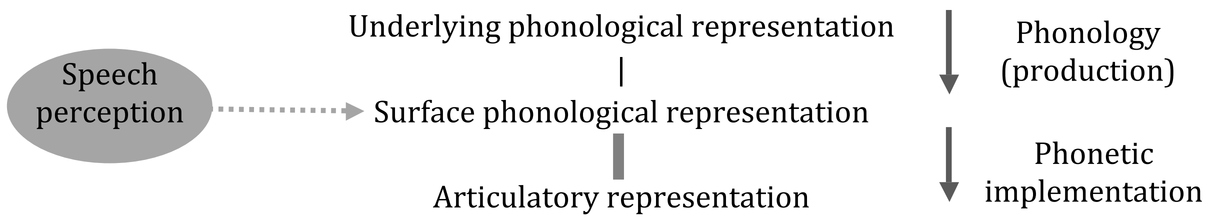

Figure 3: Three representations and the phonetics-phonology interface (thick line) as assumed in some phonological studies, with an indirect influence of speech perception (grey oval) on phonology. Processing is restricted to the production direction (arrows on the right).

In contrast to the model given in Figure 3, one could consider speech perception as the mapping of a phonetic representation onto a surface phonological representation, thus approximately the reverse of phonetic implementation. If we include the perception/recognition direction of processing in our linguistic model (as forwarded in linguistics by e.g. Pierrehumbert 1990), then the model looks as in Figure 4. In such a model, the phonetics-phonology interface is the mapping of surface onto phonetic form in the production direction and the reverse in the comprehension direction.

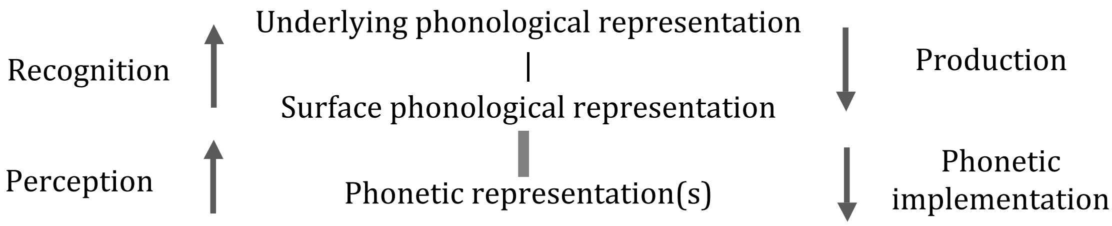

Figure 4: Two phonological representations and at least one phonetic representation, with two directions of processing (arrows left and right of the representations). The phonetics-phonology interface (bold line) depends on the number and ordering of the phonetic form(s).

In Figure 4, we see the comprehension process on the left and the production process on the right (as proposed in psycholinguistics by Cutler & Norris 1979 and McQueen & Cutler 1997, and in linguistics by Boersma 2007). In the present article, we will term the mapping of phonetic onto surface form in the comprehension direction 'perception', and the mapping of surface onto underlying form in the same direction '(phonological) recognition'. In the production direction, the mapping of underlying to surface form is '(phonological) production', and the mapping of surface to phonetic form 'phonetic implementation' (following Boersma 2007).

One phonetic representation does not seem to be sufficient in the model of Figure 4. In speech comprehension, listeners have an auditory form as input, whereas in speech production, speakers produce an articulatory form as output. It is therefore reasonable to assume that we have both an articulatory and an auditory phonetic representation. These two phonetic forms can be connected to the surface form in three different ways: First, both auditory form and articulatory form can be connected independently to the surface representation. Second, only the articulatory form can be connected to the surface representation, and the auditory form is then attached to the articulatory form. And third, only the auditory form could be connected to the surface form, and the articulatory form is then connected to the auditory form. See Figure 5 for an illustration of these three possibilities.

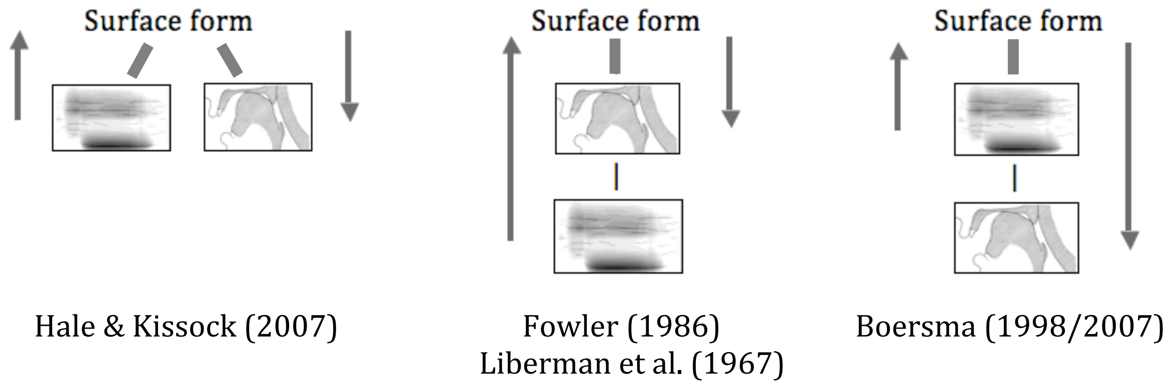

Figure 5: The three possible connections between the phonetic forms and the surface representation. The auditory representation is given as a cochleagram (of the German sound [ki], recorded and transformed with Praat: Boersma & Weenink 2009), the articulatory representation is given as an x-ray tracing (of German /k/, from Wängler 1958). The phonetics-phonology interface is

indicated by a thick line. The directions of processing are indicated by the arrows; an upward arrow for perception, a downward arrow for phonetic implementation.

All three possibilities have been proposed. Model one and three we encountered already in §3.2. Hale & Kissock (2007) attach both the auditory and the articulatory form to the surface form, as in Figure 5 on the left. In this model, the two directions of processing involve different mappings: In perception from auditory form to surface form, in phonetic impelmentation from surface form to articulatory form. Boersma (1998/2007) connects only the auditory form directly to the surface form, as in Figure 5 on the right: In perception, the listener maps the auditory form onto the surface form; in phonetic implementation, the speaker does the reverse (with the same connections), and additionally maps the auditory onto the articulatory form.[6]

Figure 5 in the middle is the reverse of Boersma's; here the articulatory form is attached to the surface representation, and this connection is forming the phonetics-phonology interface. The auditory form is connected to the surface form only indirectly, and plays only a small role in speech perception and none in articulation. This view has been forwarded by Fowler (1986) as *Direct Realist Theory of Speech Perception*, which assumes that in speech perception the speech signal is directly interpreted in terms of articulatory gestures. A similar approach can be found in the *Motor Theory of Speech Perception* by Liberman and colleagues (Liberman, Cooper, Shankweiler & Studdert-Kennedy 1967, Mattingly & Liberman 1969, Liberman & Mattingly 1985). According to this theory, we can only perceive speech sounds with the help of the gestures that we use for articulation. The interest in this theory was renewed with the discovery of mirror neurons, specifically with studies showing that there is activation in the mouth area of the motor cortex when participants listened to speech sounds (e.g. Fadiga, Craighero, Buccino & Rizzolatti 2002, Watkins, Strafella & Paus 2003). Both the Direct Realism Theory and the Motor Theory of Speech Perception are, as their names imply, models of speech perception, and do not account for phonological processes.[7]

We saw now that the shape of the phonetics-phonology interface depends on a number of theoretical decisions. First we have to decide whether we make a strict distinction between phonetic and phonological representations, that is, whether we employ a modular view of phonetics and phonology. If not, the interface is difficult to define. If

---

[6] Only in the phonetic implementation does Boersma's BiPhon model (2007) differ from his earlier work (1998), see Boersma & Hamann (2009b: 11f.) for details.

[7] The primacy of the articulatory representation advocated in these two theories encounters two problems. First, infants can perceive speech sounds much earlier than they are able to produce them (e.g., Jusczyk 1997). If speech perception depended on the availability of gestures, this would not be possible. Second, in speech production the target is assumed to be articulatory and not auditory, but bite-block experiments (Lindblom, Lubker & Gay 1979) showed that speakers can adjust to articulatory obstructions very quickly, which indicates that they most likely have an auditory rather than an articulatory target in their production process (Boersma to appear: 22). These two points are also problematic for Hale & Kissock's proposal. The assumed equal status and independency of the auditory and the articulatory representation cannot account for the primacy of perception in the acquisition process. The findings in bite-block experiments are problematic since the articulatory form is assumed to be independent of the auditory form, thus a direct correction of articulation via the perceptual representation is not possible.

we assume modularity, a further question is whether we want to include both directions of processing, speech production and speech comprehension, in our linguistic model. If we restrict ourselves to the production direction, then the interface is the connection between phonological surface forms and their articulatory correlates. If we include speech comprehension, we have to further ask ourselves what phonetic form is connected to the phonological surface representation. Here the interface can take any of the three possibilities given in Figure 5. Independent of the last two questions, we have to decide in a modular theory whether the connection between phonological surface form and phonetic representation(s) is automatic or language-specific. If it is automatic, then the phonetics-phonology interface is the same for all languages, if it is language-specific, we have to define the mapping between phonological surface form and phonetic form for every language.

In the following section, we will have a closer look at one phonological model that formalizes the interface as the language-specific mapping between surface form and auditory form. We will investigate what its underlying assumptions predict for the phonetics-phonology interface, and how these predictions relate to experimental findings in phonetics, phonology and psycholinguistics.

4 An explicit model of the phonetics-phonology interface

The *Bidirectional Phonology and Phonetics* model (Boersma 2007, 2009, to appear; henceforth: BiPhon) is a modular approach to phonetics and phonology which includes both the production and the comprehension direction. It assumes an arbitrary connection between phonemes and their phonetic form, as promoted by Keating (1988) and Pierrehumbert (1990: 377ff.). All connections between representations, both phonological and phonetic, are formalized in the same way in BiPhon, namely as ranked OT-constraints.[8] The representations and the constraints for their evaluation are given in Figure 5.[9]

[8] A formalization with weighted constraints, as in *Harmonic Grammar* (Smolensky & Legendre 2006) is also possible, see e.g. Boersma & Escudero (2008) and Boersma & Hamann (2008). The connections could in principle also be formalized in an connectionist approach, see e.g. Plaut, McClelland, Seidenberg & Patterson (1996) on the connection between graphemes and phonemes, McClelland & Rumelhart (1981) and Rumelhart & McClelland (1986) on morphophonemic alternations, and McClelland & Elman (1986) on the modeling of speech perception.

[9] The BiPhon model has been supplemented by a semantic module and respective representations, see Apoussidou (2007) and Boersma (to appear).

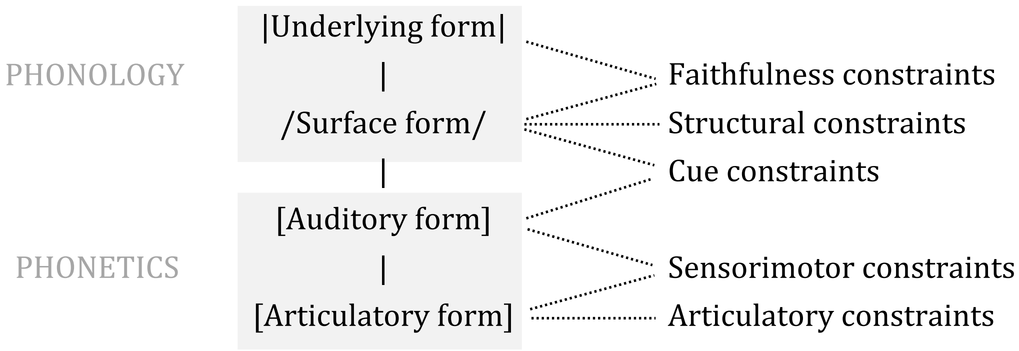

Figure 5: Representations and connections in the BiPhon model.

The faithfulness constraints that map underlying to surface form and vice versa are well known from the OT literature, where they are usually employed only in the production direction. The structural constraints replace partly the usually employed markedness constraints. In contrast to the latter, structural constraints only evaluate restrictions on the phonological surface form (such as *NoCoda, *ComplexOnset, Trochee, etc.) but do not refer to any phonetic details (such as *Retroflex or MinDist, recall the minimal functional approach described in §3.1). Restrictions on phonetic forms are formalized by cue constraints, sensorimotor constraints, and articulatory constraints.

Cue constraints are responsible for the mapping of auditory cues onto phonological surface form. They constitute the phonetics-phonology interface. The cue constraint *[low F1]/a/, for instance, expresses the fact that the auditory cue of a low first formant should not be perceived as the vowel /a/.[10] In the production direction, the same cue constraint is interpreted as the fact that the low vowel /a/ should not be realized with a low first formant. Examples for the language-specificity of cue constraints are given in §4.1 below. These cue constraints allow any arbitrary connection between a phonological category and an auditory dimension.

Sensorimotor constraints evaluate the mapping of articulation to auditory form. They "express the language user's knowledge of the relation between articulation and sound; with them, the speaker knows how to articulate a given sound and can predict what a certain articulatory gesture will sound like" (Boersma 2009: 60). The sensorimotor constraints are usually not formalized, as they are not language-specific. Articulatory constraints regulate the restrictions on the articulatory form. An example is the constraint *Distance(i-R), which stands for "avoid articulating sequences of high front vowels and retroflex sounds" (Hamann 2003: 180).

All constraints apply bidirectionally, i.e. in both directions of processing. The strict division between phonology and phonetics is not only found in the representations, but also in the constraints evaluating them, as we saw in the replacement of markedness constraints by structural constraints if they concern the phonological surface structure, and by cue or articulatory constraints if they concern phonetic representations (therefore the criticism on markedness constraints voiced by Haspelmath 2006 and Hume 2004 does not apply).

[10] Cue constraints can also refer to other phonological categories such as features, syllables, feet, etc. Syllable boundaries are e.g. part of the cue constraints in 4.1.

4.1 The language-specific phonetics-phonology interface

The BiPhon model assumes a language-specific mapping between auditory form and phonological surface representation, i.e. a non-universal interface. An example from foreign-language perception will illustrate that this assumption and its formalisation can account for observed phenomena. The example is the German perception of Dutch voiceless plosives, illustrated with the perception of the Dutch word *pitten* /.pɪ.tə./ 'to sleep'. This word is usually perceived by German listeners with no knowledge of Dutch as the German word *Bitte* /.bɪ.tə./ 'please': The initial voiceless plosive is perceived as a voiced one because it lacks the aspiration that is typical for German initial voiceless plosives. Instead of the descriptive account given in the previous sentence, BiPhon allows to formalise this observation and to make testable predictions for other word forms and the perception and production of other languages by German native speakers.

In order to be able to account for the perception of /.pɪ.tə./ as /.bɪ.tə./, we have to specify the auditory realisation of the Dutch word. The phonological form /.pɪ.tə./ has the auditory form transcribed here as [_ᵖɪt˺_ᵗə]. In this form, "_" stands for the silence of the voiceless closure phase, "ᵖ" and "ᵗ" stand for the bilabial and alveolar voiceless release bursts, "ɪ" and "ə" stand for the respective formant values of the two Dutch vowels, and "t˺" stands for the formant transitions into an alveolar stop (see Boersma & Hamann 2009b for similar transcriptions). This notation is employed to make the distinction between the abstract phonemes in /.pɪ.tə./ and the gradient auditory form that is associated with it explicit. How German listeners then perceive [_ᵖɪt˺_ᵗə] is determined by their language-specific cue constraints and the ranking of the constraints. These cue constraints connect the phonological categories that are employed in the language (in our restricted case /p/ and /b/) to the employed auditory dimension (here: aspiration noise [ʰ], silence [_], and burst noise [ᵖ ᵇ]). The cue constraints and the perception of [_ᵖɪt˺_ᵗə] by a German listener is shown in Tableau (1), where only those cue constraints are given that are relevant for the perception of the initial plosive.

(1) German perception of Dutch initial /p/, i.e. the sound [_ᵖ]

[_ᵖɪt˺_ᵗə]	*[_ᵖ]/.p/	*[_ᵖʰ]/.b/	*[_ᵖʰ]/.p/	*[_ᵖ]/.b/
/.pɪ.tə./	*!			
☞ /.bɪ.tə./				*

The cue constraint *[_ᵖ] /.p/ in Tableau (1) reads as "do not perceive the auditory form of a plain (i.e. unaspirated) labial plosive as an initial voiceless /p/", the constraint *[_ᵖʰ]/.b/ as "do not perceive the auditory form of an aspirated labial plosive as an initial voiced /b/", and so forth. The constraints and their ranking, i.e. the German perception grammar, mirror the fact that in German, plosive releases of initial voiceless labials with aspiration are usually perceived as /p/, and plosive releases of initial voiceless labials without aspiration are usually perceived as /b/.

The Dutch perception of the same auditory form results in a different output, as illustrated with Tableau (2). In the second and third cue constraint, "‿" stands for a voiced closure phase, and "^{b}" for the bilabial voiced release bursts, the cue constraint *[‿^{b}]/.p/ thus prohibits the perception of a voiced closure together with a voiced labial release as an initial voiceless /p/. The constraints and their ranking express the fact that in Dutch, plosive releases of initial voiceless labials without aspiration are usually perceived as /p/, and voiced closure phases with voiced labial releases as /b/.

(2) Dutch perception of Dutch initial /p/, i.e. the sound [_^{p}]

[_^{p}ɪt˺_^{t}ə]	*[_^{p}]/.b/	*[‿^{b}]/.p/	*[‿^{b}]/.b/	*[_^{p}]/.p/
☞ /.pɪ.tə./				*
/.bɪ.tə./	*!			

We can see that with such formalization, the differences in phonetic details between languages can be made explicit and can account for differences in cross-language speech perception that have been observed in psycholinguistic studies (see e.g. Werker & Logan 1985, Strange 1995).[11]

The same constraints and constraint ranking as in the perception processes are also used in the phonetic implementation of segments, where they can account for a foreign-language accent of an L2 speaker. If a German speaker with no knowledge of Dutch phonetics tries to produce the word *pitten*, she will use the same cue constraints and rankings as in Tableau (1), i.e. her native grammar. When used for phonetic implementation, the cue constraints have to be interpreted in the other direction, for instance, the constraint *[_^{p}] /.p/ reads as "do not realize the surface form of an initial /p/ as the auditory form [_^{p}]", and so forth. The constraints and their ranking then formalize the fact that German initial /p/ is usually realized as voiceless plosive with aspiration noise, and initial /b/ as voiceless plosive without aspiration noise. In our example, we assume that the German speaker has created a surface form /.pɪ.tən./ for *pitten*. The phonetic implementation process for German is shown in Tableau (3), for Dutch in Tableau (4), both focussing on the initial plosive.

(3) German phonetic implementation of an initial /p/

/.pɪ.tən./	*[_^{p}]/.p/	*[_^{ph}]/.b/	*[_^{ph}]/.p/	*[_^{p}]/.b/
[_^{p}ɪt˺_^{tn}]	*!			
☞ [_^{ph}ɪt˺_^{tn}]			*	

[11] The German "mis"-perception of [_^{p}] as /b/ formalized in Tableau (1) could be argued to be due to a lexical effect, since German has the word Bitte /.bɪ.tə./ 'please' but no word like Pitte with the surface form /.pɪ.tə./. That the lexicon can influence the perception was shown by Ganong (1980) and Samuel (1987). However, an identification task with four native speakers of German without any knowledge of Dutch showed that the participants, who had to identify Dutch sequences of obstruent plus vowel /a/ as native German sequences, identified Dutch [_^{p}a] as German /ba/ in 42.98% of the cases (these results stem from the control group in Hamann & Sennema 2005 but are not reported in the article).

(4) Dutch phonetic implementation of an initial /p/

/.pɪ.tən./	*[_^{p}]/.b/	*[˷^{b}]/.p/	*[˷^{b}]/.b/	*[_^{p}]/.p/
[˷^{b}ɪt˺_^{t}ə]		*!		
☞ [_^{p}ɪt˺_^{t}ə]				*

The phonetic implementation in Tableau (3) thus accounts for the typical German aspiration of the voiceless, non-aspirated initial plosives in Dutch, cf. Tableau (4).

It could be argued that the differences we saw between German and Dutch phonetic implementation and perception do not lie in the mapping from auditory onto surface form and vice versa, but in a different surface (and underlying) representation. We could for instance assume that the Dutch /b/ – /p/ contrast is one represented with the feature [±voice], whereas the German is one of the feature [±spread glottis]. This would allow for a universal phonetics-phonology interface, i.e. a language-independent filling in of the auditory cues and articulatory gestures depending on the features used in the language (see Hale & Kissock's 2007 approach in §3.2). For labial plosives, the feature [+voice] would then translate into /b/, /–voice/ and [–spread glottis] into /p/, and [+spread glottis] into /p^{h}/. Two aspects oppose such a universal interpretation. First, there are small phonetic differences between languages with the same feature specification, e.g. [+spread glottis] means a stronger aspiration in Korean than in English (Kim 1994, Kim, Beddor & Horrocks 2002), hence we would have to add some additional device to our theory to account for these differences. Second, there are phonological reasons independent of the phonetic realization for using the same phonological feature across languages. In our example, both Dutch and German /p/ are usually referred to by the phonological feature [–voice] (e.g., Booij 1995 for Dutch and Wiese 1996 for German), because in both languages /p/ groups with other voiceless obstruents. In German, /p/ forms a group with the other aspirated plosives /t/ and /k/ but also with the non-aspirated fricatives /f/, /s/ and /ʃ/ in the process of final devoicing, where voiced obstruents are devoiced in syllable-final position. The same process applies in Dutch, which lead phonologists to describe the process in both languages as a change in feature value from [+voice] to [–voice]. This identical formalization allows for an easier cross-linguistic comparison of the same or similar phonological processes. It is, however, a purely phonologically motivated use, to simplify cross-linguistic comparisons. It differs from the description with the phonetically more faithful features [±spread glottis] for German and [±voice] for Dutch.[12] In BiPhon, it is assumed that we do not have language-independent phonetic

[12] Here we see an illustration of Ladefoged's (1988) point that the use of features and their phonetic correlates, i.e. the phonetics-phonology interface in Ladefoged's view, strongly depend on the aims of the phonological description (see footnote 2): if we aim at a phonological typological description, then we should use the same phonological features for the same/similar cross-linguistic processes, and if we aim at a universal phonetic interpretability of phonological features, we should use the same phonetic correlates for the same features. We cannot achieve both at the same time.

interpretability of features, for arguments given above. This allows us to use the same phonological features for the same processes.[13]

4.2 Bidirectionality of the interface

In BiPhon, all constraints are used bidirectionally, i.e. both in the comprehension and the production direction. We saw above that cue constraints form the phonetics-phonology interface in BiPhon. How do they behave in the two processing directions?

In perception, cue constraints interact with structural constraints. This means that speech perception is not only influenced by our native-language categories via the cue constraints (which map auditory information to phonemes or phonological features), but also by structural restrictions such as syllable structure and phonotactics. Psycholinguistic studies have shown that listeners cannot perceive a difference between words that deviate minimally from their native phonotactics and the corresponding correct words. This is illustrated with the well-known example of Japanese listeners who are not able to discriminate between [ebzo] and [ebɯzo] (Dupoux, Kakehi, Hirose, Pallier & Mehler 1999). In the BiPhon model, this case of phonological knowledge influencing the perception is formalized as a structural constraint against coda consonants (*CODA, simplifying the Japanese facts slightly) being higher ranked than the cue constraint that militates against the interpretation of nothing in the auditory form as a high back vowel (*[ ]/u/).[14]

The formalization of perception as interaction of cue and structural constraints in BiPhon can furthermore account for the fact that loanword adaptation via perception often yields other outputs than we would expect based on native phonological processes. Boersma & Hamann (2009b) show that this is an automatic outcome of the model, and does not require loanword specific devices or constraint rankings as proposed in other studies (e.g. Rose & Demuth 2006; Shinohara 2004).

In the production direction, the use of cue constraints can explain why phonological surface structures can influence the phonetic realizations. This has been observed for instance in prosodic strengthening, where initial segments are increasingly "strengthened" the higher they occur in the prosodic hierarchy. Strengthening can be realized in several ways such as lengthening, more amplitude, and more linguo-palatal contact (see e.g. Fougeron 2001; Wightman, Shattuck-Hufnagel, Ostendorf & Price 1992). In BiPhon, such strengthening is the result of position-dependent cue constraints, which are necessary anyway to account for e.g. syllable-dependent allophonic realizations.

In sum, the formalization of assumptions in BiPhon makes it possible to compare the outcome of this model with experimental data. Formalization can thus help us getting a grip at the phonetics-phonology interface and test the validity of the assumptions made in a model.

---

[13] BiPhon is neutral with respect to the question whether phonological features are universal or not.

[14] This constraint can only be low ranked in Japanese because the high back vowel /u/ is very often devoiced in Japanese and thus has often inaudible cues. Cue constraints are hence not low ranked in order to explain the observed phenomena, but because the cues they refer to are not salient. High ranked cue constraints, on the other hand, mirror the fact that the cues they refer to are salient and important for the listeners of the respective language.

5 Conclusion: No theory-neutral interface

The present article looked at several proposals on the phonetics-phonology interface. We saw that the definition of the interface depends strongly on our phonological assumptions. Do we want to give up on an abstract phonological form at all, as in Exemplar Theory? If not, do we want to employ a modular view of phonetics and phonology, i.e., make a strict distinction between phonetic and phonological representations? If modularity is assumed, is perception part of the linguistic model or not? If perception is included, how is the auditory form integrated in the model? The scientist choosing a modular theory has to further decide whether the connection between phonological surface form and phonetic representation is automatic or language-specific. All of these choices result in a different phonetics-phonology interface.

In §4 we looked in detail at one proposal, namely the BiPhon model, with an explicit formalisation of a modular theory that includes both directions of processing and works with language-specific mappings between phonological and phonetic form. We saw that this model makes predictions that can be checked against experimental data. Instead of arguing for a specific model, the aim of this section was to illustrate the need for explicit models and formalisations. A linguistic model that formalizes assumptions and therefore provides us with testable predictions is preferable to approaches that simply describe observations. It is obvious that we need further models with explicit formalizations. Only alternative models based on different assumptions of phonology and its consequences for the phonetics-phonology interface can help us decide which phonological assumptions can best account for the observed data. Phonologists and phoneticians interested in the phonetics-phonology interface need to think in terms of models that can formalize the interface. These models, in turn, have to be informed again by experimental findings. As Fromkin (1975) has written:

> [T]he study of phonology cannot be divorced from phonetics if phonologists aim to provide explanatory and predictive theories of sound systems, rather than merely descriptive theories. (p. 104)

References

Apoussidou, Diana (2007). *The learnability of metrical phonology*. Ph.D. dissertation, University of Amsterdam.

Archangeli, Diana and Douglas Pulleyblank (1994). *Grounded Phonology*. Cambridge (Mass.): MIT Press.

Beckman, Mary (2003). Input representations (Inside the mind and out). In: G. Garding and M. Tsuijmura (eds.) *Proceedings of the 22nd West-Coast Conference on Formal Linguistics*. Somerville, MA: Cascadilla Press, 70-94.

Blevins, Juliette (2004). *Evolutionary Phonology*. Cambridge: Cambridge University Press.

Boersma, Paul (1998). *Functional phonology: formalizing the interaction between articulatory and perceptual drives*. Ph.D. dissertation, University of Amsterdam.

Boersma, Paul (2007). Some listener-oriented accounts of *h-aspiré* in French. *Lingua* 117: 1989-2054.

Boersma, Paul (2009). Cue constraints and their interactions in phonological perception and production. In: Paul Boersma and Silke Hamann (eds.) *Phonology in perception*. Berlin: Mouton de Gruyter, 55-108.

Boersma, Paul (to appear). A programme for bidirectional phonology and phonetics and their acquisition and evolution. In: Anton Benz and Jason Mattausch (eds.) *Bidirectional Optimality Theory*.

Boersma, Paul and Paola Escudero (2008). Learning to perceive a smaller L2 vowel inventory: an Optimality Theory account. In: Peter Avery, B. Elan Dresher and Keren Rice (eds.) *Contrast in phonology: Theory, perception, acquisition*. Berlin: Mouton de Gruyter, 271-301.

Boersma, Paul and Silke Hamann (2008). The evolution of auditory dispersion in bidirectional constraint grammars. *Phonology* 25: 217-270.

Boersma, Paul and Silke Hamann (2009a). Introduction: models of phonology in perception. In: Paul Boersma and Silke Hamann (eds.) *Phonology in Perception*. Berlin: Mouton, 1-24.

Boersma, Paul and Silke Hamann (2009b). Loanword adaptation as first-language phonological perception. In: Leo Wetzels and Andrea Calabrese (eds.) *Studies in loan phonology*. Amsterdam: Benjamins, 11–53.

Boersma, Paul and David Weenink (2009). *Praat: doing phonetics by computer (version 5.1.04) [Computer program]*. Manuscript, Retrieved April 08, 2009 from http://www.praat.org/.

Booij, Geert (1995). *The Phonology of Dutch*. Oxford: Oxford University Press.

Broselow, Ellen (2004). Language contact phonology: richness of the stimulus, poverty of the base. *Proceedings of the North-Eastern Linguistic Society* 34: 1-22.

Chomsky, Noam and Morris Halle (1968). *The Sound Pattern of English*. New York: Harper and Row.

Clements, George N. (1985). The geometry of phonological features. *Phonology Yearbook* 2: 225–252.

Cohn, Abigail C. (1993). Nasalization in English: phonology or phonetics? *Phonology* 10: 43-81.

Cohn, Abigail C. (2007). Phonetics in Phonology and Phonology in Phonetics. *Working Papers of the Cornell Phonetics Laboratory* 16: 1-31.

Cutler, Anne and Dennis Norris (1979). Monitoring sentence comprehension. In: William E. Cooper and Edward C.T. Walker (eds.) *Sentence processing: psycholinguistic studies presented to Merrill Garrett*. Hillsdale: Erlbaum Associates, 113-134.

Donegan, Patricia and David Stampe (1979). The study of Natural Phonology. In: Dinnsen (eds.), 126-173.

Dupoux, Emmanuel, Kazuhiko Kakehi, Yuki Hirose, Christophe Pallier and Jacques Mehler (1999). Epenthetic vowels in Japanese: a perceptual illusion. *Journal of Experimental Psychology: Human Perception and Performance* 25: 1568-1578.

Fadiga, Luciano, Laila Craighero, Giovanni Buccino and Giacomo Rizzolatti (2002). Speech listening specifically modulates the excitability of tongue muscles: A TMS study. *European Journal of Neuroscience* 15: 399-402.

Flemming, Edward (1995/2002). *Auditory Representations in Phonology*. London: Routledge.

Flemming, Edward (1997). Phonetic detail in phonology: towards a unified account of assimilation and coarticulation. In: K. Suzuki and D. Elzinga (eds.) *Proceedings of the 1995 Southwestern Workshop in Optimality Theory (SWOT)*. University of Arizona.

Flemming, Edward (2001). Scalar and categorical phenomena in a unified model of phonetics and phonology. *Phonology* 18: 7-44.

Flemming, Edward (2003). The relationship between coronal place and vowel backness. *Phonology* 20: 335–373.

Flemming, Edward (2004). Contrast and Perceptual Distinctiveness. In: Bruce Hayes, Robert Kirchner and Donca Steriade (eds.) *The Phonetic Bases of Markedness*. Cambridge: Cambridge University Press, 232-276.

Flemming, Edward (2005). Speech perception and phonological contrast. In: David B. Pisoni and Robert E. Remez (eds.) *The Handbook of Speech Perception*. Oxford: Blackwell, 156-181.

Fougeron, Cécile (2001). Articulatory properties of initial segments in several prosodic constituents in French. *Journal of Phonetics* 29: 109-135.

Fowler, Carol A. (1986). An event approach to the study of speech perception. *Journal of Phonetics* 14: 3-24.

Fromkin, Victoria (1975). The interface between phonetics and phonology. *Proceedings of the 8th International Congress of Phonetic Sciences*, Leeds.
Ganong, William F. (1980). Phonetic categorization in auditory word perception. *Journal of Experimental Psychology: Human Perception and Performance* 6: 110-125.
Gordon, Matthew (2007). Functionalism in phonology. In: Paul de Lacy (eds.) *The Cambridge Handbook of Phonology*. Cambridge: Cambridge University Press, 61-77.
Hale, Mark (2000). Marshallese phonology, the phonetics-phonology interface and historical linguistics. *The Linguistic Review* 17: 241-257.
Hale, Mark and Madelyn Kissock (2007). The phonetics-phonology interface and the acquisition of perseverant underspecification. In: Gillian Ramchand and Charles Reiss (eds.) *The Oxford Handbook of Linguistic Interfaces*. Oxford: Oxford University Press, 81-102.
Hall, T. A. and Silke Hamann (to appear). On the cross-linguistic avoidance of rhotic plus high front vocoid sequences. *Lingua*.
Hamann, Silke (2003). *The Phonetics and Phonology of Retroflexes*. Ph.D. dissertation, Utrecht Institute of Linguistics.
Hamann, Silke and Anke Sennema (2005). Voiced labiodental fricatives or glides - all the same to Germans? In: Valery Hazan and Paul Iverson (eds.) *Proceedings of the Conference on Plasticity in Speech Processing*. London: UCL, 164-167.
Harris, John and Geoff Lindsey (1995). The elements of phonological representation. In: Jacques Durand and Francis Katamba (eds.) *Frontiers of Phonology: Atoms, Structures, Derivations*. Harlow, Essex: Longman, 34-79.
Haspelmath, Martin (1999). Optimality and diachronic adaptation. *Zeitschrift für Sprachwissenschaft* 18: 180-205.
Haspelmath, Martin (2006). Against markedness (and what to replace it with). *Journal of Linguistics* 42: 25-70.
Hume, Elizabeth (2004). Deconstructing markedness: a predictability-based approach. *Proceedings of the Berkeley Linguistics Society* 30: ?
Hume, Elizabeth and Keith Johnson (2001). A model of the interplay of speech perception and phonology. In: Elizabeth Hume and Keith Johnson (eds.) *The Role of Speech Perception in Phonology*: Academic Press, 3-26.
Hyman, Larry M. (2001). The limits of phonetic determinism in Phonology: *NC revised. In: Elizabeth Hume and Keith Johnson (eds.) *The Role of Speech Perception in Phonology*. San Diego: Academic Press, 141-185.
Jakobson, Roman, Gunnar Fant and Morris Halle (1952). *Preliminaries to Speech Analysis: The Distinctive Features and their Correlates*. Cambridge (Mass.): MIT Press.
Johnson, Keith (1997). Speech perception without speaker normalization: An exemplar model. In: Keith Johnson and John W. Mullennix (eds.) *Talker variability in speech processing*. New York: Academic Press, 143-165.
Johnson, Keith (2005). Decisions and Mechanisms in Exemplar-based Phonology. *UC Berkeley Phonology Lab Annual Report* 2005: 289-311.
Jusczyk, Peter W. (1997). *The Discovery of Spoken Language*. Cambridge (Massachusetts): MIT Press.
Keating, Patricia A. (1988). The phonology-phonetics interface. In: Frederick Newmeyer (eds.) *Linguistics: the Cambridge survey. Volume I: Linguistic Theory: Foundations*. Cambridge: Cambridge University Press, 281-302.
Kenstowicz, Michael (1994). *Phonology in Generative Grammar*. Cambridge: Blackwell.
Kim, Mi-Ran Cho (1994). *Acoustic characteristics of Korean stops and perception of English stop consonants*. Ph.D. dissertation, University of Wisconsin, Madison.
Kim, Mi-Ryoung, Patrice Speeter Beddor and Julie Horrocks (2002). The contribution of consonantal and vocalic information to the perception of Korean initial stops. *Journal of Phonetics* 30: 77-100.
Kingston, John (2006). The phonetics-phonology interface. In: Paul de Lacy (eds.) *The Cambridge Handbook of Phonology*. Cambridge: Cambridge University Press, 401-434.
Kingston, John and Mary Beckman (1990). *Papers in Laboratory Phonology I: Between the Grammar and Physics of Speech*. Cambridge: Cambridge University Press.
Kirchner, Robert (1997). Contrastiveness and faithfulness. *Phonology* 14: 83-111.

Kirchner, Robert (1998/2001). *An Effort-Based Approach to Consonant Lenition*. New York: Routledge.

Kirchner, Robert (1999/2004). Phonological contrast and articulatory effort. In: Bruce Hayes, Robert Kirchner and Donca Steriade (eds.) *Phonetic Basis of Markedness*. Cambridge: CUP, 25 p.

Ladefoged, Peter (1972). Phonetic prerequisites for a distinctive feature theory. In: A. Valdman (eds.) *Papers in linguistics and phonetics to the memory of Pierre Delattre*. The Hague: Mouton, 273-286.

Ladefoged, Peter (1975). *A Course in Phonetics*. New York: Harcourt Brace Jovanovich.

Ladefoged, Peter (1988). The many interfaces between phonetics and phonology. *UCLA Working Papers in Phonetics* 70: 13-23.

Liberman, A. M., F. S. Cooper, D. P. Shankweiler and M. Studdert-Kennedy (1967). Perception of the speech code. *Psychological Review* 74: 431-461.

Liberman, Alvin M. and I.G. Mattingly (1985). The motor theory of speech perception revised. *Cognition* 21: 1-36.

Lindblom, Björn (1986). Phonetic universals in vowel systems. In: John J. Ohala and Jeri J. Jaeger (eds.) *Experimental phonology*. Orlando: Academic Press, 13-44.

Lindblom, Björn, James Lubker and Thomas Gay (1979). Formant frequencies of some fixed-mandible vowels and a model of speech motor programming by predictive simulation. *Journal of Phonetics* 7: 147-161.

Lisker, Leigh and Arthur S. Abramson (1964). A cross-language study of voicing in initial stops: acoustical measurements. *Word* 20: 384-422.

Lisker, Leigh and Arthur S. Abramson (1970). The voicing dimension: Some experiments in comparative phonetics. *Proceedings of the 6th International Congress of Phonetic Sciences*, Prague, Academia Publishing House: 563-567.

Mattingly, I.G. and Alvin M. Liberman (1969). The speech code and the physiology of language. In: K.N. Leibovic (eds.) *Information processing in the nervous system*. New York: Springer.

McCarthy, John J. and Alan Prince (1993). *Prosodic morphology I: Constraint interaction and satisfaction*. Manuscript, University of Massachusetts and Rutgers University.

McClelland, James and Jeffrey L. Elman (1986). The TRACE Model of Speech Perception. *Cognitive Psychology* 18: 1-86.

McClelland, James and David D.E. Rumelhart (1981). An interactive activation model of context effects in letter perception: Part 1. An account of basic findings. *Psychological Review* 88: 375-407.

McQueen, James and Anne Cutler (1997). Cognitive processes in speech perception. In: William J. Hardcastle and John Laver (eds.) *The handbook of phonetic sciences*. Oxford: Blackwell, 566-586.

Mielke, Jeff (2004). *The emergence of distinctive features*. Ph.D. dissertation, Ohio State University.

Ohala, John J. (1981). The listener as a source of sound change. In: Carrie Masek, Roberta A. Hendrick and Mary Frances Miller (eds.) *Proceedings of the Chicago Linguistics Society 17*. Chicago, 178-203.

Ohala, John J. (1983). The origin of sound patterns in vocal tract constraints. In: P.F. MacNeilage (eds.) *The production of speech*. New York: Springer, 189-216.

Ohala, John J. (1992). Alternatives to the sonority hierarchy for explaining segmental sequential constraints. *Proceedings of the Chicago Linguistic Society* 26: 319–338.

Ohala, John J. (2005). Phonetic explanations for sound patterns: Implications for grammars of competence. In: William J. Hardcastle and M. Beck (eds.) *A figure of speech: A festschrift for John Laver*. London: Erlbaum, 23-38.

Padgett, Jaye (2001). Contrast dispersion and Russian palatalization. In: Elizabeth Hume and Keith Johnson (eds.) *The Role of Speech Perception in Phonology*. San Diego (CA): Academic Press, 187-218.

Padgett, Jaye (2003a). Contrast and post-velar fronting in Russian. *Natural Language and Linguistic Theory* 21: 39-87.

Padgett, Jaye (2003b). The emergence of contrastive palatalization in Russian. In: Eric Holt (eds.) *Optimality Theory and Language Change*. Dordrecht: Kluwer Academic Press, 307-335.
Passy, Paul (1891). *Etude sur les changements phonétiques et leurs caractères généraux*. Paris: Librairie Firmin-Didot.
Pater, Joe (2004). Bridging the gap between perception and production with minimally violable constraints. In: René Kager, Joe Pater and Wim Zonneveld (eds.) *Constraints in phonological acquisition*. Cambridge: Cambridge University Press, 219-244.
Pierrehumbert, Janet B. (1990). Phonological and phonetic representation. *Journal of Phonetics* 28: 375-394.
Pierrehumbert, Janet B. (2001). Exemplar dynamics: word frequency, lenition, and contrast. In: Joan Bybee and Paul J. Hopper (eds.) *Frequency Effects and the Emergence of Linguistic Structure*. Amsterdam: Benjamins, 137-157.
Plaut, David C., James McClelland, Mark S. Seidenberg and Karalyn Patterson (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review* 103: 56-115.
Port, Robert and Michael L. O'Dell (1985). Neutralization of syllable final voicing in German. *Journal of Phonetics* 13: 455-471.
Prince, Alan and Paul Smolensky (1993 [2004]). *Optimality Theory: Constraint Interaction in Generative Grammar*. London: Blackwell.
Rose, Yvan and Katherine Demuth (2006). Vowel epenthesis in loanword adaptation: representational and phonetic considerations. *Lingua* 116: 1112-1139.
Rumelhart, David D.E. and James McClelland (1986). On learning the past tenses of English verbs. In: David D.E. Rumelhart, James McClelland and the PDP Research Group (eds.) *Parallel Distributed Processing. Volume 2*. Cambridge (Massachusetts): MIT Press, 216-271.
Samuel, Arthur G. (1987). Lexical uniqueness efects on phonemic restoration. *Journal of Memory and Language* 26.
Shinohara, Shigeko (2004). Emergence of Universal Grammar in foreign word adaptations. In: René Kager, Joe Pater and Wim Zonneveld (eds.) *Constraints in phonological acquisition*. Cambridge: Cambridge University Press, 292-320.
Smolensky, Paul and Géraldine Legendre (2006). *The Harmonic Mind*. Cambridge (Massachusetts): MIT Press.
Stampe, David (1972). *How I speant my summer vacation*. Ph.D. dissertation, University of Chicago.
Steriade, Donca (1995). *Positional neutralization*. Manuscript, University of California.
Steriade, Donca (1997). *Phonetics in phonology: the case of laryngeal neutralization*. Manuscript, University of California.
Steriade, Donca (1999). Licensing laryngeal features. *UCLA Working Papers in Phonetics* 3: 25-146.
Steriade, Donca (2001a). Directional asymmetries in place assimilation: a perceptual account. In: Elizabeth Hume and Keith Johnson (eds.) *The Role of Speech Perception in Phonology*. San Diego (CA): Academic Press, 219–250.
Steriade, Donca (2001b). *The phonology of perceptability effects: the P-map and its consequences for constraint organization*. Manuscript, University of California.
Strange, Winifred (1995). Cross-language studies of speech perception: a historical review. In: Winifred Strange (eds.) *Speech perception and linguistic experience: Issues in cross-language research*. Timonium: York Press, 3-45.
Wängler, Hans H. (1958). *Atlas deutscher Sprachlaute*. Berlin: Akademie-Verlag.
Warner, Natasha, Allard Jongman, Joan Sereno and Rachel Kemps (2004). Incomplete neutralization and other sub-phonemic durational differences in production and perception: evidence from Dutch. *Journal of Phonetics* 32: 251-276.
Watkins, Kate E., A.P. Strafella and Tomáš Paus (2003). Seeing and hearing speech excites the motor system involved in speech production. *Neuropsychologia* 41: 989-994.
Werker, Janet F. and John S. Logan (1985). Cross-language evidence for three factors in speech perception. *Perception and Psychophysics* 37: 35-44.
Wiese, Richard (1996). *Phonology of German*. Oxford: Oxford University Press.

Wightman, C.W., Stefanie Shattuck-Hufnagel, M. Ostendorf and Patti J. Price (1992). Segmental durations in the vicinity of prosodic phrase boundaries. *Journal of the Acoustical Society of America* 91: 1707-1717.

Wright, Richard (2001). Perceptual Cues in Contrast Maintenance. In: Elizabeth Hume and Keith Johnson (eds.) *The Role of Speech Perception in Phonology*. San Diego: Academic Press, 251-277.