Chapter 6
Phonological systems

Harry van der Hulst

We define phonology as a study of the perceptible side of language, of both speech and sign. Phonologists (and perhaps speakers, too) use abstract symbolic representations to cope with the endless variability in actual utterances. This chapter focuses on the phonology of spoken languages. The symbolic representations are meant to capture phonetic properties that are distinctive. Representations are compositional, consisting of pivotal units called phonemes, smaller units (features), and larger units (syllables). Given a distinction between symbolic units (features, phonemes etc.) and the actual perceptible signal, there must be a system which brings the symbolic representations closer to articulations and to acoustic (or visual) events. Consequently, phonological analyses of languages typically start out with a phonetic representation, which in addition to the distinctive properties of phonemes also contains predictable properties, called allophonic. Phonetic representations are reduced to phonemic representations using criteria such as complementary distribution and phonetic similarity between allophones; then phonotactic analysis determines the inventory of phonemes in terms of constraints on feature combinations and phoneme combinations. The combining of morphemes into complex words may result in violations of the usual constraints on phoneme combinations. For example, in English, a combination of the morphemes *in* and *possible* causes the /n/ to be replaced by /m/ (*impossible*). Such rules lead to allomorphy, i.e. phonemic form variation in morphemes. However, it is shown that not all allomorphy is driven by phonotactic constraints. This chapter draws attention to different points of view, especially in relation to the phenomenon of neutralization.

Keywords: phonemic level, phonetic level, phoneme, implementation, neutralization, allophony, allomorphy.

1. Introduction

In this chapter I present a 'conservative' view of phonology, i.e. a view that recognizes the phoneme/allophone distinction and a distinction between three rules types (morpho-lexical rules, phonemic rules and allophonic rules). The model attributes a central role to constraints (both at the phonemic and the phonetic level), not as the sole agents in phonology (as in

Optimality Theory[1]), but rather as driving the application of (repair) rules (again at both levels). A model of this sort captures what in my view are unavoidable distinctions, although many models have tried to eliminate aspects of it. Along the way, I will indicate how various models deviate from this conservative model. Section 2, 3 and 4 address some general, preliminary issues, paving the way for section 5 which discuss the notions *levels* and *rules*. Section 6 offers some concluding remarks which make brief reference to several important topics that could not be covered in this chapter.

2. Is phonology necessary?

I take it to be self-evident that language, as a collection of utterances, is dependent on a grammar, which is a property of human brains. This **mental grammar** is a system of basic units and constraints on their combinations which generates (or recognizes) an infinite set of linguistic expressions that have a meaning and a form. A perhaps surprising recent debate among linguists is whether the form side of linguistic expressions includes information about **perceptible events** (e.g. events that can be heard or seen when linguistic expressions are realized as utterances). In the current view of Noam Chomsky linguistic expressions are purely mind-internal (mental states), their function being 'to organize thought'. In this view, the mental grammar is primarily a **syntactic module** that generates these expressions in terms of a syntactic structure which groups basic units (packages of a conceptual structure and a syntactic labelling) into labeled constituents. In this view, the 'form' of basic units and expressions *is* the syntactic labelling and the syntactic structure with labelled terminal and non-terminal nodes (the **syntactic form**). This syntactic module, with the infinite set of expressions that it characterizes, is called **I(nternal)-language**. I-language, in Chomsky's view, does not contain a module that deals with the perceptible utterances which, for him, fall in the domain of **E(xternal)-language**. There must thus be an **externalization module** which relates the internal expressions to external perceptible events. This externalization is brought about by linking the concept/label packages referred to above to a third type of information which specifies the perceptible events (as psycho-acoustic templates and as articulatory plans). This information is also usually called 'form', in this case meaning **perceptible form**. Here I define phonology as the module that achieves the externalization of linguistic expressions. It seems to me that this module is just as internal as Chomsky's notion of I-language in the sense of being a mental system. The claim that the externalization system does not belong to the so-called 'narrow language faculty' is a purely terminological matter,

---

[1]   See Prince and Smolensky (1993/2004), McCarthy (2001).

especially since it is far from obvious that the syntactic module comprises mechanisms that are unique to language any more or less than the externalization module.

For our purposes it is noteworthy that Chomsky's conception of the mental grammar delegates one aspect of utterances that is traditionally seen as the domain of syntactic form to phonological form, namely the linear order in which the minimal concept-label packages occur once they have been externalized in perceptible form. The need for linear order is, as the argument goes, entirely caused by the external perceptible properties which, given their nature, must mostly be produced (and perceived) one after the other.[2] Thus, Chomsky's syntactic structures are like mobiles and it is the task of 'phonology' to serialize the terminal elements of this structure. In this chapter I will not deal with this kind of external 'syntactic phonology', which I happily will leave to the syntactians. Clearly, even though the *necessity* for linearization can be attributed to the nature of the externalization system (and the perceptual systems that process utterances back to linguistic expressions), the principles or conventions that govern the *specific* linearizations that languages display are mostly non-phonological.

3. Phonology = phonemics and phonetics

To both lay people and many linguists it is not clear why there are two disciplines that deal with the perceptible form of linguistic expressions. Traditionally, both **phonology** and **phonetics** seem to refer to the study of sound. We must first note that the use of the morpheme 'phon' is a historical accident, due to the fact that linguists did not until recently include sign languages in their studies. This leads to the often heard equation 'language = speech', which in addition to conflating internal and external language, implies that the external side always uses the auditory channel. If we include sign languages, the 'phon' part of these terms is an anachronism.[3] For this reason, I will consistently speak of the perceptible form of language, rather than the audible or visible form. Any humanly perceptible form would in principle qualify as the potential form side of language. But why two disciplines? While phonology and phonetics are different (according to some, completely non-

---

[2]   In support of accounting for linear order in the phonology, we note that sign languages allow a higher degree of parallelism of the basic packages because the visual system can compute independent articulatory actions of the hands, the face and the body simultaneously. Parallelism also occurs in spoken language, for example with reference to the parallelism of intonation and segmental properties of words, or in case of so-called (e.g. tonal) surprafixes.

[3]   See Lillo-Martin and Sandler (2006) for a broad overview of the linguistic study of sign language.

overlapping) fields, in the tradition of **American Stucturalism**, phonology is a term that includes phonetics and, in addition, **phonemics**. In this view, phonetics studies the complete array of physical aspects of language form (in terms of acoustic, auditory and articulatory properties), while phonemics focuses on properties of the signal that are **distinctive** (or **contrastive**), i.e. that distinguish one **morpheme** from another. Thus, phonetics covers both distinctive and predictable properties (which are not treated differently at this level), in practice (but not in principle) abstracting away from properties that are due to **paralinguistic factors**, i.e. physical and psychological properties of the speaker. Phonemics, on the other hand, studies the way in which morphemes and words are stored in the **mental lexicon**, on the assumption that at this *level* only distinctive properties are specified. In practice (but not in principle) non-distinctive properties are then assumed to be **un(der)specified** in the lexical representations.

Attention for the phonemic angle became prevalent toward the end of the 19th century when the development of increasingly sophisticated phonetic equipment brought to light the apparent infinite variability of the speech signal and of the '**speech sounds**' 'in' it. To explain the obvious fact that language users neglect most of that variability when they 'home in' on the linguistic expression, it was proposed that the perceptible form of this expression must be represented in terms of units that lie beneath all this 'irrelevant' variability.[4] This step can be identified as the reason for distinguishing between **phonemes** (relevant, i.e. distinctive properties) and 'speech sounds' (or **allophones**; 'all' properties) and thus between phonemics and phonetics.[5] In mostly European schools, the term 'phonology' came to be used for 'phonemics'. Here we follow the American Structuralists and use the term phonology in the general sense of covering both phonemics and phonetics.

---

[4]   The development of the International Phonetic Alphabet (IPA), despite its use of the term 'phonetic', was in fact meant have a set of symbols for these units that were stripped of their 'irrelevant' properties. The justification for calling this system 'phonetic' is that in the initial stage of analysis, the linguists wants to keep his or her options open as to which phonetic properties are going to be 'relevant' or not.

[5]   The 'reduction' from variable perceptual inputs to abstract mental 'concepts' is an essential cognitive instrument for dealing with the infinite variability of percepts, not just in processing language utterances, but in processing all information. From this perspective, phonemic representations are 'conceptual structures' for parts of speech events and phonemes are 'central concepts' (although features would be the basic concept); see Taylor (2003) for a cognitive view on phonology that emphasizes the role of categorization.

4. The general form of a phonological theory

As shown in Anderson (1985), phonological theorizing (like many other human systems of thought) seems to follow a dialectic pattern such that the attention and creativity of scholars shift back and forth between what are the two major aspects of any phonological system: **representations** (involving the identification of units and their combinations at various levels) and **derivations** (dealing with **relations** between **levels**). In a general sense, then, all phonological theories agree that phonology is an input-output mapping system:

(1) **F** (Input): output

The function **F** forms the **derivational aspect** of the theory, whereas both inputs and outputs form the **representational aspect**. The derivational aspect deals with the issue of how many levels are distinguished and how these levels are related. The representational aspect concerns the structure of each level in terms of its basic units, their grouping and the constraints that capture their combinations. As shown in (2), it is possible that the phonological module contains more than one F function, such as an **allophonic function** (F1) and an **implementation function** (F2)

(2) **F1** (phonemic representation): allophonic/phonetic representation

    Phonology

    F2 (phonetic representation): phonetic event

It is generally assumed that representations have a discrete compositional structure, which means that there are **basic units** (**primes**), which are either deemed universal and innate or language-specific and learned (or a combination of these), that are combined in accordance with universal and language-specific **constraints**. A major development in phonology in the 1940s, due to the work of Roman Jakobson, has been that the original primes, i.e. the *sequential* phonemes (taken thus far to be indivisible, holistic units), were 'decomposed' into smaller *simultaneous* units, called (**distinctive) features** (Jakobson, Fant and Halle 1952). Motivation for such smaller sub-phonemic units was that rules often refer to what appear to be non-arbitrary (i.e. recurrent) groupings of phonemes. We see this, for example, in the English aspiration rule which accounts for the aspirated allophones of the *group* /p, t, k/ in the onset of stressed syllables (see section 4). This grouping can be found in other rules or constraints in English or other languages. Assuming that the members of this group share a property, or set of properties, which are represented in terms of features, the aspiration rule can make reference to these features rather than to a (random) list of phonemes. With

features, the change from /p/ to [pʰ] does not involve the substitution of one 'whole segment' by another, but rather can be stated as the change or addition of specific property.[6]

In the remainder of this chapter we focus on derivational issues. For reasons of space, there is no review of representational issues.

## 5. Derivational issues: levels and rules

5.1 Three levels

The recognition of phonemes as distinct from allophones implies that we recognize at least two **levels of representation**: the phonemic level and the phonetic (or allophonic) level. The phonemic representation can be identified as the manner in which the phonological form of linguistic expressions is represented in the mental lexicon. However, the phonetic representation, although closer to the actual perceptible event, is usually not identified with it. Most phonologists take both the phonemic and the phonetic level to be mental and **symbolic**. At the phonetic level, the symbols (i.e. features) refer to phonetic events (articulatory plans or actions and psycho-acoustic 'images' or actual acoustic properties). We could regard the phonetic level as an **articulatory plan**: "The result of phonological encoding [of the lexical form] is a *phonetic* or *articulatory plan*. It is not yet overt speech; it is an internal representation of how the planned utterance should be articulated – a program for articulation" (Levelt 1989: 12). Whether the features at the phonological level also refer to phonetic events has remained a matter of dispute and in some approaches, phonemic symbols are taken to be **substance-free**, their only role being 'to be different' so that they can differentiate morphemes (see Hjemslev 1953). A more common view is to regard phonological features (taken to characterize units at both the phonemic and phonetic level) as **mental concepts** of phonetic events. (Below I discuss the view that the phonetic level requires additional phonetic features or distinctions.)

That the phonetic representation is symbolic is suggested by the use of IPA symbols for units at this level, which imply a **discreteness** that is not present in the physical (articulatory or acoustic) event. Indeed, the term 'speech sound', as a 'slice' of the speech signal, refers to an abstract unit. Thus, if one refers to IPA symbols at the phonetic level as 'speech sounds',

---

[6] In this case the IPA practice to represent aspiration in terms of a superscripted diacritic symbols anticipated a feature analysis. In general, IPA diacritics can be set to be 'featural'. But apart from that IPA treats phonemes as holistic. Bell's visible speech transcription system was almost entirely feature-based (Bell 1867).

this is misleading since the term refers to symbolic units which, as such, are not present in the speech signal. This means that there is yet a third 'level' which a phonological model must recognize, which is the output of a system of **phonetic implementation**, i.e. a set of rules that count as instructions to a device (such as the human vocalization system) that will deliver a physical event that is based on the articulatory plan (i.e. the phonetic representation). We thus end up with the following phonological model:

(3)   Phonemic representation (lexicon)

↓

Allophonic rules

↓

Phonetic representation (articulatory plan)

↓

Implementation

↓

Physical event (utterances)
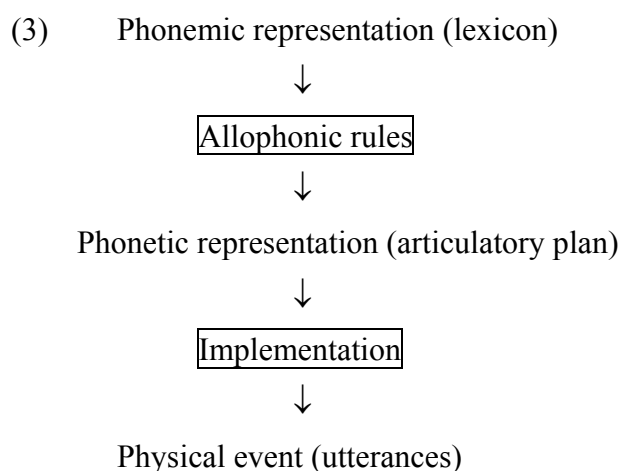
Most phonologists would assume that not only the implementation system but also the allophonic system apply 'on line', i.e. in the actual externalization of a linguistic expression. In this chapter, I will not discuss phonetic implementation explicitly.

Allophonic properties can depend on the position of phonemes with respect to the edge of constituents (such as syllable, foot or the word). The allophonic rules that account for them are 'habits of articulation' (Hockett 1955) that characterize the speech characteristics of a specific language. These rules are not universal (not all languages aspirate stops), although they are phonetically 'natural' or phonetically 'motivated', i.e. grounded in phonetic 'tendencies'. These rules are often said to be **automatic** in that they apply wherever the appropriate context is met. In this sense, such rules *cannot* have arbitrary exceptions.[7]

The rather simple diagram in (3) raises a host of questions, answers to which lead to different phonological theories. Theories differ in terms of how many levels are taken to be

---

[7]   Here we must reckon with the fact that allophonic processes may constitute an innovation in the language and as such be subject to lexical diffusion, i.e. gradual integration in the lexicon, targeting certain classes of words (e.g. the more frequent ones) before others. The end point of lexical diffusion can leave some lexemes unaffected, which, then, are technically exceptions. New allophonic processes can also lead to variability. Harris (1987) has shown that allophonic rules can sometimes display sensitivity to morphological structure.

necessary to connect the representation that language users store in their mental lexicon (i.e. the phonemic representation) to physical events as well as in the way that 'comparable levels' are defined (see section 5.6.2). Additionally, if rules are ordered (e.g. within the allophonic function), there will be so-called **intermediate levels**. However, even though there are several different views on the nature of these levels and about procedural methods to derive phonemes from phonetic transcriptions, it would seem that a distinction between a phonemic and a phonetic/allophonic level has been foundational to the development of phonology.[8]

At the level of language description, however, the phoneme/allophone distinction is 'universally' accepted. It is taught to all students who take linguistic courses and forms the basis of all language description. In practical terms, Pike (1948) refers to phonemics as a technique to reduce speech to writing, since writing systems predominantly provide symbols for distinctive units, and indeed, most often use alphabetic units that symbolize phonemes.[9] In this context it is important to see, however, that the 'emic/etic' distinction is not committed to slicing the speech signal into the kinds of units that are represented in alphabetic writing systems. If one were to hold the view that the basic distinctive units are bigger (e.g. syllable-sized) or smaller (features), the same distinction can be applied. Correspondingly, writing systems can be based on such larger or smaller units (see Rogers 2005 and chapter 4, this volume).

5.2 The phonetic level

An important (and unresolved) question is what kind of variation is taken to fall within the realm of allophony (as opposed to what is referred to the implementation system). Even when one excludes variation due to paralinguistic factors and restrict oneself to phonetic properties that are due to the linguistic environment (i.e. surrounding phonemes, stress and word edges) language descriptions (i.e. IPA transcriptions) often ignore 'very minute' variation, which is then referred to as resulting from **co-articulation** when the variation is due to the influence of neighboring phonemes, but even positionally determined properties may be relegated to another 'lower' level of analysis (i.e. implementation). For example, in phonological works on English, the slight difference in place of articulation for the phoneme /k/ in words like

---

[8]   See van der Hulst (2013) for historical overviews and references.

[9]   It forms the basis for the earliest phonographic writing systems. One of the earliest works on the development of an alphabetic writing system by the First Grammarian embodies an explicit account of the phoneme–allophone distinction. This shows that this distinction was recognized long before the dawn of phonology in linguistics.

cool and *keel* is taken for granted. Implicitly or explicitly, this minute variation is relegated to phonetic implementation, implying that 'proper allophony' is part of the phonology. In practice, we see that allophonic rules are mostly proposed for predictable properties that could have been distinctive, as witnessed by the fact in some other language they are distinctive. For example, aspiration is distinctive in Thai where minimal pairs can be found (as in 4a) that contrast only in this difference, which leads the phonological analyses in (4b) for English and Thai, respectively:

(4) a. [pɔ:n] to wish [pʰɔ:n] also

[tɔp] to support [tʰɔp] be suffocated

[kat] to cut [kʰat] to polish

b. English Thai

/p/ /pʰ/ /p/

[pʰ] [p] [pʰ] [p] Allophones
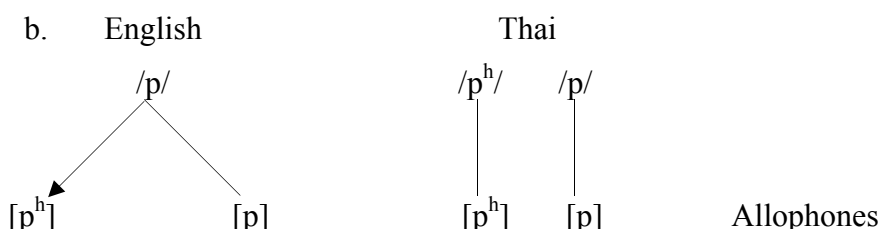
Knowing what kinds of properties could be distinctive, but not yet knowing that they actually are distinctive in the language at issue, prompts the linguist to take note of them in the IPA transcription. However, limiting allophonic variation to variation that could be distinctive is *not* an official rule of phonological analysis. Actual descriptive practices differ and many descriptions will also write allophonic rules for phonetic properties that have not been found to be ever distinctive. In fact, IPA, with its rich set of **diacritics** allows for a **narrow transcription** which captures minute co-articulatory or positional effects (although up the limit that is allowed by the diacritics).[10] Arguably, a **broad transcription** could be defined as one that aims at registering potentially distinctive properties. But this practice is not always followed. For example, many analyses of English propose an allophonic rule that derives an unreleased variant of stops in English (e.g. [p̚]) in syllable final position. However, a difference between released and unreleased stops has never been shown to be distinctive in any language and there is thus no generally recognized distinctive feature that allows us to express this difference. An additional issue is that one cannot always be sure

---

[10] See Pullum and Ladusaw (1986) for a comprehensive survey of commonly used phonetic symbols. In the study of phonological acquisition by children or of phonological disorders it may be crucial to record very fine phonetic detail, irrespective of the issue of potential distinctiveness and we see that for that purpose special IPA-version with many additional diacritics have been developed.

whether a case of phonetic variation should be expressed in terms of a distinctive feature because it is not clear how to value a phonetic distinction as falling within the scope of a given distinctive feature. Distinctive features do not necessary correlate with the exact same phonetic property (or set of properties) cross-linguistically.

The view of a phonetic representation as encoding predictable properties that are, at least potentially, distinctive is thus not the only view, although that view is necessary if one adopts the strategy that the phonemic and the allophonic level are stated in terms of one set of features. However, another view of what ought to be expressed at the phonetic level is that this level should encode everything that is under the 'control' of the speaker which means that it regards phonetic effects that are characteristic of a specific language and thus not universally present in all languages as a matter of phonetic necessity. In this view everything that goes above and beyond the call of what is universally necessary (given the nature and working of articulatory organs) belongs to the phonetic level (see Anderson and Lightfoot 2011 for a defence of this view). This view implies that the vocabulary of the phonetic level is different from that of the phonemic level. It requires a special set of phonetic features in addition to the distinctive features. In addition, Anderson and Lightfoot (2011 chapter 6) argue that what must be expressed at the phonetic level is not just what can be stated in positive properties, it can also regard negative properties when a language displays a suppression of some phonetic tendency, e.g. the suppression of phonetic pitch raising after voiceless consonants in languages that use pitch distinctively. The difference between the two views that we have discussed is dependent on whether non-universal phonetic processes must be attributed to the phonetic level (keeping implementation universal) or not; in the latter case we allow phonetic implementation to include language-specific properties (see Pierrehumbert 1990).

One possible conclusion (which gives credit to both views) is that it may be necessary to recognize that specific cases of phonetic variation are ambiguous, i.e. analyzable as allophonic or as implementational, given that the implementational system is not limited to universal effects only. This, I would argue, is not a bad result, because ambiguity of this kind will allow phonological processes to migration from one submodule (implementation) to the next (allophony), and from there on to the phonemic level (see section 5.6).

5.3 The phonemic level: redundant properties

Allophonic rules specify contextually (i.e. syntagmatically) predictable properties of phones at the phonetic level. In this section, we discuss the idea that phonemes *as such* have

predictable properties as well. We must try to establish whether these predictable properties are truly different from those that have been called allophonic.

For example, one commonly finds the claim that in a typical 5 vowel system (/i/, /u/, /e/, /o/, /a/), the value for the feature [±round] is fully predictable in terms of the following rules, where the value of α is either + or −:

(5) a. [+low] → [-round]
     b. [-low, αback] → [αround]

In Chomsky and Halle (1968) (henceforth: SPE, for *The Sound Pattern of English*) rules of this sort are called **segment structure rules (**or **conditions**). Predictable specifications of this kind are often called **redundant** and for this reason rules like (5) are also called **redundancy rules**.

However, the context-free segment structure rules are not the only rules that characterize redundancy in phonemic representations. Let us consider an example of a context-sensitive redundancy. In English we can find words that start with three consonants (*splash, strap, squash*). In such cases the initial consonants is always /s/, the second is always a stop and the third is always an approximant. This means that the phonemic string /spl/ or /str/ can be represented with minimal specification, leaving out many feature specifications of /s/, /p/ or /t/ and /l/ or /r/. In fact the first consonant need not be featurally specified at all. All of its properties are predictable. For the second and third consonant we only need to know that it is a labial and approximant, respectively

(6) /s/   /p/   /l/
      -     [–continuant]   [+approximant]
            [+labial]

It would that seem that, just as there are **paradigmatic constraints** on the combination of features within a phoneme (as expressed in segment structure rules or conditions), there are also **syntagmatic constraints** on feature specifications. These constraints are often called **phonotactic constraints** (which in some uses also includes the segment structure rules). In the case of (6) many feature specifications are dependent on the mere presence of other consonants in the word-initial cluster. While these, then, are predictable, they do not fall within the scope of allophony since there is no complementary distribution of two phonetic entities. The 's' phone in (6) is not in complementary distribution with any other phonetically similar phone. Rather, we are dealing here with specific statements about the **defective distribution** of phonemes. In the spot before 'pl' no other consonant than /s/ can appear. To

put it differently: in this position all consonantal distinctions are **neutralized** and the product of this neutralization is /s/. We see a similar distributional fact in languages that limit the number of consonants that can occur in the syllabic coda position (Itô 1986). The fact is that there are constraints regarding the distribution of phonemes and such constraints entail predictabilities regarding the feature specification of consonants that occupy positions that are excluded for other consonants. These predictabilities hold at the phonemic level and are thus different in nature from the syntagmatic predictabilities that hold at the phonetic level and which are captured in allophonic rules.

Whether redundancies that follow from the context-free, paradigmatic constraints should be stated at the phonemic level or can be included in the set of allophonic rules is not so easy to decide. It could be argued that the predictable (redundant) roundness of back vowels falls within the same realm as the predictable (allophonic) aspiration of voiceless obstruents in English. In both cases, one might argue that the predictable properties have the function of enhancing the phonetic identity of phonemes and that, in some other language, they could be distinctive. For example in Turkish the roundness value of /u/ is not predictable because this language has a back unrounded vowel at the same vowel height. We can then say that phonemes can have predictable properties which occur in all their occurrences or in occurrences in certain positions and there is no obvious reason for why we would not capture both in allophonic rules. However, if there has to be a level with fully specified phonemic units (stripped only of contextual predictabilities), it would have to follow that paradigmatic constraints must be met at the phonemic level in which they cannot be collapses with contextual allophonic rules.

Concluding then that fully specified phonemes contain redundant properties (at the very least those that are syntagmatically conditioned), it has been assumed in several models that redundant properties at the phonemic level remain *unspecified* in the lexical representation. This requires redundancy statements to trigger rules that fill in redundant values at some point in the derivation, presumably before phonetic implementation and perhaps even before the application of allophonic rules. SPE decided against such underspecification on technical grounds. Redundancies would still be captured in rules or conditions that function as indications of the 'low cost' of redundant specifications. An early argument in favor of underspecification based on 'saving memory space' is no longer deemed convincing given the vast storage capacity of the human mind. How could we decide between the full-specification theory and the impoverished theory? A possible window on how the human brain represents the form of words lexically comes from research that uses neuro-imaging

techniques (see Lahiri and Reetz 2010). At present it is still in dispute what this kind of work is going to show.[11]

We have thus far assumed that in case of contrast, where both values of a feature are non-redundant, both values of a feature must be lexically specified. This is called **contrastive specification**. In Archangeli (1984) and Kiparsky (1982) it is proposed that even in case of contrast one of the values can be left unspecified. This increases the degree of underspecification since now, in addition to leaving out truly redundant values, we also leave out one value for each distinctive feature. This is called **radical underspecification**. In this approach, we get an additional set of 'fill-in' rules (often called **default rules**) which specify for each feature the **complement value** of the **lexically-specified value**. For example, in a language having a contrast between stops and fricatives we could specify only the latter as [+continuant], which requires a default fill-in rule:

(7) [∅continuant] → [+continuant]

A decision on which value is left unspecified would have to be based on the 'activity' of this value in the sense of playing an active part in the phonology. The default value could be universally fixed or what is default could be language-specific. Archangeli (1984) suggests that there is a universal bias for the default, but allows that languages may overturn this bias. The default value is also often called the **unmarked value**, which seems fitting, because it is literally unmarked. This would make the lexically specified value the **marked value**.

Unary features (also called privative, single-valued) embody a radicalization of radical underspecification in the sense that it is now claimed that for each distinctive dimensions only one 'value' plays a role in the phonology. This 'value' then is equal to the feature label itself, [round], [nasal], [low], etc. Anderson and Ewen's (1987) **Dependency Phonology** model presents the most explicit defense of unary features. Clearly, a unary feature theory does not have default fill-in rules. **Government Phonology** (Kaye, Lowenstamm and Vergnaud 1985) claims that the unary features all have stand-alone potential, i.e. each of them by themselves represents a possible phoneme. There can still be redundancy in a model of this kind in the sense that for any given phoneme system certain unary features may be predictable on the basis of others. For example, a feature [ATR] (Advanced Tongue Root)

---

[11]   Even in a full specification approach it is not the case that each segment must be specified for every feature. We need to know which features are *applicable* to different types of phonemes. It can be argued that features like [continuant], [lateral], [anterior] etc. are not ever needed for vowels. Likewise there are features that we need for vowels, but never for consonants.

would be redundant for high vowels if there is no ATR-distinction at this height, and assuming that the high vowels would be phonetically [ATR]. It is even possible to conceive of radical underspecification, if a system of this sort would make use of the null-option. For example, a three vowel system (/i/, /u/, /a/) could be represented as follows:

(8) /i/   /u/   /a/

   -   U   A   (Default element: I)

For extensive discussion of binary and unary feature theories, see Ewen and van der Hulst (2000).

In conclusion, we have two understandings of the phonemic level, one impoverished, being stripped of redundant values and one fully specified. Dresher (2009) offers a recent defense of the minimal, i.e. impoverished view, on the grounds that this approach explains why redundant value would appear to play no active role in phonological rules that apply at the phonemic level (i.e. the rules we discuss in section 5.5).

5.4 Phonemic overlap and the morphophonemic level

Returning to the allophonic rule of English that creates aspirated variants for voiceless stops (see (4)) we note that in this case the phonemes /p/, /t/ and /k/ each have two allophones that uniquely belong to each phoneme. Voiceless stops occur in aspirated and unaspirated variants which are in **complementary distribution** (which means that they occur in non-overlapping sets of contexts). The standard analysis has been to analyze aspiration as contextually conditioned, occurring before a stressed vowel (unless preceded by /s/), with unaspirated stops occurring *elsewhere*. This then leads to postulating the **phonemes** /p/, /t/ and /k/ (where the symbol choice reflects that the unaspirated stop is seen as 'the basic variant') and an aspiration rule (see (4b)) which accounts for the predictable property of aspiration, creating aspirated **allophones** [pʰ], [tʰ] and [kʰ] in addition to the allophones [p], [t] and [k] (which reflect the 'direct realization' of the phoneme). However, it may also occur that an allophone counts as the manifestation of two different phonemes such as the flap [ɾ] in American English which is an allophone of the /t/ and the /d/ when these two phonemes occur intervocalically with stress on the preceding vowel, as in [raɪɾəɹ], which is the phonetic form for 'writer' and 'rider'[12]. This is known as **phonemic overlap**:

---

[12]   I ignore here the possibility of a difference in the length of the vowel, due to the difference between a following /d/ and /t/ at the phonemic level.

(9)   /t/   /d/   Phonemic level

                              Allophonic rules

     [t]   [tʰ]   [ɾ]   **[d]**   Allophonic level
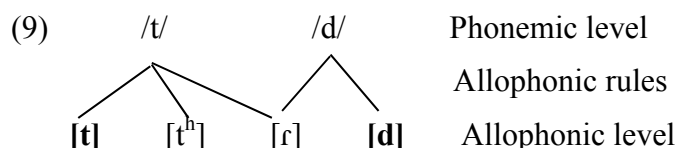
(There may of course be additional allophones for both phonemes; the bold allophones are taken to be the 'basic' ones.)

Rules that create phonemic overlap are called **neutralizing rules**. Neutralization occurs when in a given context the contrast between two phonemes of language L is neutralized.[13] In the case of flapping the *product of neutralizing* the contrast between /t/ and /d/ is a shared phone that is unique to that context. In other cases, the phone can be identical to another allophone of one of the phonemes. A famous example of this occurs in Dutch where in syllable final position there is no contrast between voiced and voiceless obstruents. The product of neutralization is a voiceless obstruent. Consider the following pairs of words:

(10)   [hɔnt]   'dog' SG   [hɔndən]   'dog' PL

   [wɑnt]   'wall' SG   [wɑndən]   'wall' PL

The suffix *-ən* indicates plurality. Observe that the final [t] corresponds to [d] when the plural suffix is present. This means that [t] **alternates with** [d]. We can analyze this by postulating that the morphemes in question end in /d/ in their lexical representation (as witnessed by the plural) and that this /d/ gets 'realized' as [t] when it occurs word finally. The observed alternation thus provides evidence for the allophonic rule of **final devoicing** which responds to the phonetic constraint in (11a):[14]

(11) a. * [-son, +voice])$_\sigma$

     b. [d] → [t] / _)$_\sigma$

The effect of this rule is that it causes phonemic overlap:

---

[13]   Neutralization of contrast can be the result of rules, such as the flapping rule, or result from defective distribution of phonemes as discussed in section 5.3. See Silveman (2012) for a broad overview of the notion neutralization.

[14]   In (4b) I informally depicted allophony as resulting from a rules that converts a phoneme into an allophone. In (11) I switch to the explicit perspective that the allophonic rules applies at the phonetic level, responding to a phonotactic constraint at this level.

(12)　　/t/　　　　/d/

[t]　　　[t]　←　[d]

(11b)
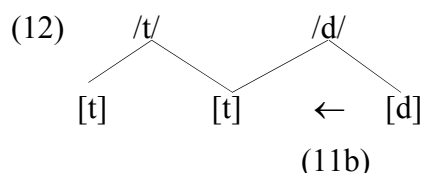
Thus, flapping and final devoicing neutralize a contrast that exist in the language under analysis. I take both rules to be allophonic rules because they are fully automatic. There are no lexical exceptions to either rule.

We should note that final devoicing, just like flapping and aspiration, creates what I will call **phonetic allomorphy**, i.e., allomorphy due to an allophonic rule.

(13) Aspiration: for 'invite ~ invit-ee': [ɪnvaɪt] ~ [ɪnvaɪt^h]

Flapping: for 'write ~ writ-er': [raɪt] ~ [raɪɾəɹ]; for 'ride' ~ 'rid-er'; [raɪd] ~ [raɪɾ]

Final devoicing: for 'hond ~ hond-en': [hɔnt] ~ [hɔnd]

The analysis of flapping and final devoicing as allophonic rules departs from the American Structuralist school which adopted a principle (called **biuniquess** in Chomsky and Halle 1965) stating that each phone could only be an allophone of one phoneme. In the Dutch devoicing case this implies that since [t] is clearly an allophone of /t/ (in all non-final positions as well as in final position where there is no 'd~t' alternation), it must be an allophone of /t/ in *all* final positions, even where it alternates with [d].

The biuniqueness principle disallowed American structuralists from analyzing the final devoicing alternation as rule that mediates between the phonemic and the allophonic level and, as a result, this alternation could not be captured at all. To this end they postulated an extra phonemic level, called the **morphophonemic level** (where the units, called **morphophonemes**, are placed between double slant lines) and a rule like final devoicing (now called a **morphophonemic rule**) would relate these two levels:

(14) //b//　//d//　//g//　//v//　//z//　//ɣ//　morphophonemic level

　　↓　　↓　　↓　　↓　　↓　　↓　　morphonemic rules

　/p/　/t/　/d/　/f/　/s/　/χ/　phonemic level

　｜　　｜　　｜　　｜　　｜　　｜　　(allophonic rules)

　[p]　[t]　[d]　[f]　[s]　[x]　phonetic level

The morphophonemic level represents all morphemes in an invariant form (i.e. /hɔnd/ for the two allomorphs), but American structuralists did not attribute a specific (psychologically realistic) value to this level. They saw it as a convenient way to capture alternations between what they regarded as independent phonemes.

Generative phonologists rejected the distinction between the morphophonemic and the phonemic level, following Halle (1959) who argued that this approach leads to undesirable consequences because sometimes what looks like a single process can have non-neutralizing effects in some cases and neutralizing effects in others. Consider the following case in Dutch. Voiceless stops become voiced before voiced stops:

(15) o[p] + [d]oen   >   o[bd]oen   'up+do: to put up'

   ui[t]+[b]rengen >   ui[db]rengen   'out+bring: to bring out'

   za[k] + [d]oek   >   za[gd]oek   'pocket+cloth: handkerchief'

In the first two cases we get neutralization because /b/ and /d/ are independent phonemes, distinct from /p/ and /t/. Hence changing /p/ in to 'b' and /t/ into 'd' is neutralizing. Therefore, in the morphophonemic approach described the process changes //p// and //t// into /b/ and /d/. However in the third case the rule is non-neutralizing because Dutch does not have a phoneme /g/. So, now we have to describe this as an allophonic rule which 'realizes' /k/ as [g]. It follows that we have to state what seems to be the same process twice, once as a morphophonemic rule and once as an allophonic rule:

(16) //p//   //t//   //k//   morphophonemic level

     ↓   ↓ | morphophonemic rule

   /b/   /d/   /k/   phonemic level

     | | ↓   allophonic rule

   [b]   [d]   [g]   phonetic level

Clearly, this is not a desirable result. If one has to state the same process twice, it feels like one misses a generalization. Halle (1959) argued that we therefore need to abandon the phonemic level and go directly from the morphonemic level (taken to be the lexical or underlying level) to the phonetic (surface) level. This claim went into history as implying that the phoneme was no longer needed. This, of course, is not the correct conclusion. Rather, what cases of this sort show it that we must allow allophonic process to be neutralizing (as I in fact assumed thus far), so that we can state all three changes as an allophonic rule. This implies a rejected of biuniqueness. Now the phonemic and morphophonemic level can be collapses into one level, which we should call the phonemic level (whether impoverished or fully specified) because it represents the distinctive units that are stripped of their allophonic properties.

5.5   Allomorphy

5.5.1   Morpho-lexical rules

Allophonic rules, as we have seen, create allophonic variation. In doing so they can also create what I called phonetic allomorphy, i.e. form variation of morphemes at the systematic phonetic level, such as [ɪnvaɪtʰ] in *invitee* and [ɪnvaɪt] in *invite*. We concluded that phonetic allomorphy results from rules that apply at the phonetic level, called allophony rules. We now turn to another kind of rules that apply at the phonemic level and that change the phonemic make-up of morphemes, depending on the occurrence of these morphemes in combination with other morphemes. There are in fact two major kinds of phonemic rules, namely **morpho-lexical rules** and **phonotactically motivated rules**. In this section I discuss the former class. The second class forms the subject of section 5.5.2.

The morpho-lexical form a somewhat heterogeneous group and finer distinctions have been made. Different theories and text books use these or other terms sometimes with differences in definition. SPE refers to a (not clearly defined) category of 'adjustment rules'. These are rules that "express properties of lexical formatives in restricted syntactic contexts and they modify syntactically generated structures in a variety of ways" (SPE: 236). In other words, readjustment rules bridge the morphosyntactically generated structure and the phonological component for the purpose of phonological rules that apply word-internally and across word boundaries. Those that perform segmental changes fail to qualify as 'phonological rules proper' because they are restricted to specific morphemes and/or triggered by specific morphemes. Here I will include adjustment rules within the class of morpho-lexical rules which in addition contains rules that SPE *would* treat as proper phonological rules because they *appear* to be largely phonologically conditioned, although often in addition being dependent on idiosyncratic non-phonological information such as lexical marking (for triggering or undergoing the rule) up to reference to specific morpheme and perhaps also word class. For example, in SPE a rule is proposed (called *velar softening*) that replaces the /k/ by /s/ in words like *electric ~ electricity*. There could be doubt, however, that this substitution is conditioned by phonemic context and it has been argued that the trigger here is a specific suffix, or rather a set of suffixes (see Strauss 1982).

Morpho-lexical rules create allomorphy that is clearly not due to allophonic processes. However, they share with neutralizing allophonic rules the fact that they have phonemes as outputs, which makes them **structure-preserving** (Kiparsky 1982). In the American Structuralist approach these rules, if stated at all, are included in the class of

morphophonemic rules, together with the neutralizing allophonic rules (see section 5.3.), since the biuniqueness principle forces *all* rules that seem to have phonemes as outputs in that category. In doing so this approach collapsed into one group very different kinds of rules into the morphophonemic set, namely neutralizing allophonic rules and morpho-lexical rules. SPE, on the other hand, recognizes only two levels and thus only one set of rules that also lumps together two very different kinds of rules, namely morpho-lexical rules and all allophonic rules. This aspect of standard Generative Phonology was an important reason for having to acknowledge **extrinsic rule ordering** since, as one might expect, morpho-lexical rules (here understood as applying *at* the phonemic level) tend to precede allophonic rules which apply at the phonetic level.

On a very large scale the phonological research community responded to SPE by re-establishing the distinction between morpho-lexical rules and allophonic rules in one form or another (although the rejection of biuniqueness was not disputed), leading to models in which morpho-lexical rules apply in a separate component, after or in conjunction with the morphology, and before all the allophonic rules.

Models differ with regard to the characterization of morpho-lexical rules. In some models, such as **Natural Generative Phonology** (Hooper 1986), allomorphy rules are not considered to be rules at all. Rather it is assumed that allomorphy is dealt with by listing the allomorphs in the lexicon (adopting a proposal by Hudson 1974). This could be done in a maximally economic way as follows:
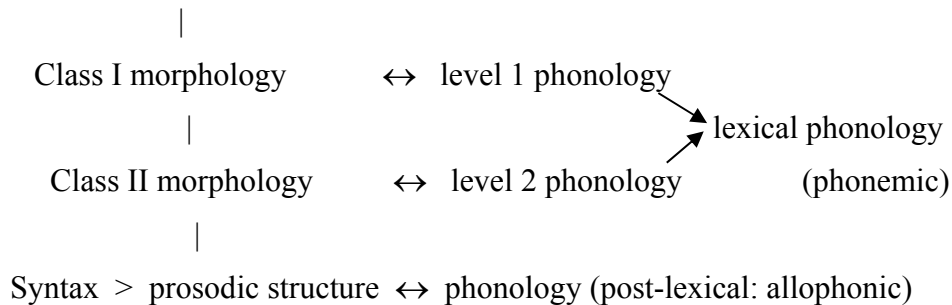
(17) /ilɛktrɪ{k/s}/

This requires a **distribution rule** which states that /s/ occurs before a specified set of suffixes. Alternatively, one could list allomorphy separately with insertion frames:

(18) /ilɛktrɪs/ [_ ity, …]
     /ilɛktrɪk/ (elsewhere)

Hudson's proposal narrows the gap between phonological allomorphy and **suppletion**. Full suppletion would occur when the two allomorphs do not share a phonemic substring and are thus listed as completely independent forms (such as /go/ and /wen/). In other proposals morpho-lexical rules are recognized *as rules*, but denied phonological status and treated as adjuncts of morphological rules.

In the lexical phonology model of Kiparsky (Kiparsky 1982), the difference between morpho-lexical rules and allophonic rules was acknowledged by relegating these classes of rules to different components of the grammar, leading to the following model:

(19)   Lexical representations

             |

Class I morphology   ↔ level 1 phonology

             | lexical phonology

Class II morphology   ↔ level 2 phonology   (phonemic)

             |

Syntax > prosodic structure ↔ phonology (post-lexical: allophonic)

Post-lexical rules are not sensitive to morphological structure or morpheme identity; they can only see phonological information. In fact, they cannot see the syntactic structure either, but rather refer to the prosodic organization that is derived from it (Nespor and Vogel 1986). By separating affixes over two levels, all rules that apply at level 1 can be formulated without much specific reference to morpheme identity and rule features so that they appear as purely phonological rules. While this is true to a large extend for a subclass of level 1 rules (namely the phonotactically motivated rules that we focus on in the next section), many instances of level 1 rule lack full generality and therefore squarely falls within the morpho-lexical rule class. Rather than 'indexing' every rule individually for those affixations that trigger them, level 1 rules are thus indexes as a group; for extensive discussion see also Straus (1982). Clearly, level ordering is a mechanism that avoids having to limit a certain rule to a subclass of affixes, which gives such rules the appearance of rules that are morphology-insensitive. The lexical model has been criticized on various grounds.

5.5.2   *Phonotactically-motivated phonemic rules*

There is a class of rules that the American Structuralist included in their class of morphophonemic rules that is much less dependent on morphological information than morpho-lexical rules. Recall that, following the principle of biuniqueness, *all* rules that output phonemes had to be included in this class. However, apart from neutralizing allophonic rules, there are two different kinds of non-allophonic rules that output phonemes. On the one hand there are the morpho-lexical rules that were discussed in the previous section, while on the other there are rules that would appear to be driven by phonotactic constraints that hold of the phonemic level. There are constraints on the combination of phonemes that refer to such domains as the syllable, the foot or the word. These constraints

capture regularities in the composition of words that need to be stated. Such constraints, called phonotactic constraints in section 5.3, effectively account for the distribution of phonemes, which is always defective, implying that no phoneme can occur in combinations (following or preceding) with all other phonemes.[15]

It is not so obvious that the /k/~/s/ rule is motivated by a phonemic phonotactic constraint that is valid for English words:

(20) * kɪ (unstressed)

If one were to pronounce *electricity* with a /k/ instead of /s/, this does not seem violate a phonotactic constraint on phoneme combinations. In other cases, such as the rule that changes the nature of vowels in *serene ~ serenity* (called 'trisyllabic laxing') we actually find exceptions such as *obese ~ obesity, nice ~ nicety* which can be taken as an indication that this change is not driven by a phonotactic constraint that holds of the word domain at the phonemic level.

Another case of allomorphy that might more plausible be said to be enforced by a phonotactic constraint involves the negative prefix in English. Consider the following word pairs:

(21) edible   in-edible
     possible   im-possible
     credible   iŋ-credible
     legal   il-legal
     regular   ir-regular

We note that the prefix form varies in its consonant: it ends in /n/, /m/, /ŋ/, /l/ or /r/. One of the constraints on phoneme sequences within simplex words in English forbids sequences like /np/ and /nk/. I will state these constraints simply as follows:

(22) * /np/, */nk/

To support the validity of these constraints, we can point to the fact that there are no monomorphemic words that display sequences of this sort. We have /pump/, but not /punp/, /iŋk/ but not /ink/. It turns out that the constraints that are valid for monomorphemic words

---

[15]   The reader should bear in mind that the term **phonotactic constraint** is relevant both with respect to the phonemic level and the phonetic level. There are phonotactic constraints that, simply put, account for wellformedness at any given level.

are also valid for a subclass of polymorphic words, i.e. domains that simplex as far as the phonology is concerned, even though they may have morphological structure. This means that when we combine the morphemes 'in' and 'possible' or 'credible' the grammar produces a unit that is ill-formed in terms of its phonemic form:

(23) i[n-p]ossible

i[n-k]redible

The idea here is that when the morphology has produced a new form, this form needs to be checked by the phonology against the list of phonotactic constraints at the phonemic level. At this point it could have been that these derivations are simply blocked, deemed to be ungrammatical. The grammar tells you: find some other way to say that something is not possible, or think it, but do not try to say it all! But this is not what happens in this case. Rather, the phonology makes available a rule to 'repair' the ill-formed product which in this case is done by replacing /n/ by either /m/ or /ŋ/.

(24) /n/→/m/

/n/→/ŋ/

We then assume that a repair rule will only come into action when a constraint violation is detected and also that there is a general way of guaranteeing that correct repair is applied. Clearly, the repair rules are not random. Before a /p/, a so-called nasal consonant can only be /m/ and before /k/ it can only be /ŋ/. Thus, given that the basic form has a nasal consonant to begin with, the repair rules make a *minimal adjustment* to achieve that the complex word no longer violates the constraints that 'flag' it as ill-formed. The motivation for taking the form with /n/ as basic is that this is the consonant that occurs when the base starts with a vowel.[16]

Note that the change in *illegal* and *irregular* does not follow from the constraints in (23) but it is likely that English has a phonotactic constraint that forbids /nl/ and /nr/, since a search in this chapter *only* deliver examples other than 'only'. We can thus posit two other repair rules:
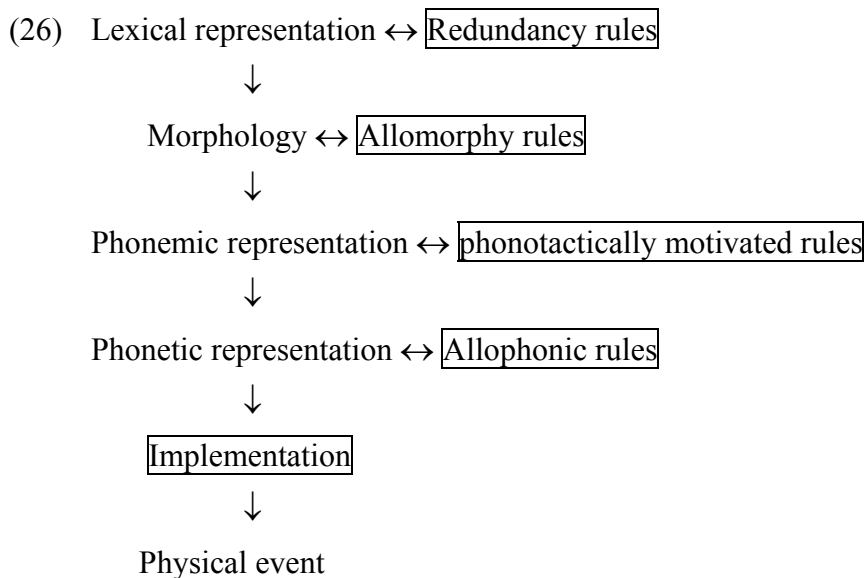
(25) /n/ → /l/

/n/ → /r/

---

[16]   There is more to this example, however, since in the word *infantile*, we replace the /n/ by a sound that does not occur as a phoneme of English, namely [ɱ]. I take this to be the result of an allophonic or implementation rule.

While it could be argued that morpho-lexical rules, if they apply actively, replace whole phonemes by other phonemes, **phonotactically motivated** rules lend themselves more readily to formulations that refer to features.

Many phonotactically-motivated rules directly serve syllable structure constraints. For example, when a language prohibits hiatus, and morphology creates a sequence /…v + v…/, repair rules must be applied. It is possible that either a consonant is inserted or a vowel is deleted. The famous case of Yawelmani, analyzed in Kisseberth (1973), who drew attention to the central role of constraints, provides a good example of insertion and vowel shortening rules that conspire to prevent syllable rhymes that consist of three phonemes (either VCC, which trigger epenthesis, or VVC which triggers vowel shortening).

Concluding, we end up with the following model:

(26) Lexical representation ↔ Redundancy rules
                    ↓
        Morphology ↔ Allomorphy rules
                    ↓
    Phonemic representation ↔ phonotactically motivated rules
                    ↓
    Phonetic representation ↔ Allophonic rules
                    ↓
        Implementation
                    ↓
        Physical event

In a model of this kind, there is very little if any need for extrinsic rule ordering, since it imposes an extrinsic ordering on phonological generalizations of different kinds. Anderson (1975), while admitting that this ordering scheme obtains in general, presents some cases in which allophonic rules seem to precede rules lower in the ranking, especially when the lower ranking rules are reduplicative operations. Many rules are not so easy to classify. Also, given that phonotactically motivated rules *can* have exceptions there is a continuum to morpho-lexical rules once exception accumulate.

As mentioned in section 1, models differ greatly with respect to the issue of levels. In some models, no sharp distinction is made between allophonic rules and the implementation module (Liberman and Pierrehumbert 1984). This is also essentially the model that is proposed in Linell (1979) who offers a very detailed and still very relevant discussion of

many aspect of phonological theory. In such one level models, one could recognize rules that apply at the phonemic level. However, some single-level models do not differentiate between phonemic rules and allophonic rules, such as **Declarative Phonology** (Scobbie 1997) and **Government Phonology** (Kaye, Lowenstamm and Vergnaud 1985, 1990). A rather extreme monostratal view (called **exemplar-based phonology;** Bybee 2001, Pierrehumbert 2001) is that lexical representations are phonetic in nature, i.e. episodic 'recordings' of actual speech events. Still other models, instead of having fewer levels, distinguish additional levels, such as for example **stratificational theory**, versions of which postulate a multitude of levels (Lamb 1966). Goldsmith (1993) proposes a model with three levels (morpheme level, word level and phonetic level) with rules applying at these levels and rules relating levels. Koskenniemi (1983) develops a two-level model with rules that states simultaneous **correspondences** between units at two levels (one lexical, the other 'surface'; this model, like SPE does also not differentiate between phonemic and allophonic relations). Optimality Theory (Prince and Smolensky 1993/2004; McCarthy 2001) was originally claimed to be a non-derivational theory. What this means is that its derivational aspect relates the lexical level to a (phonemic or phonetic) 'surface level' without the 'intermediate levels' that SPE creates due to the fact that its rules applied in an ordered fashion. By rejecting rule ordering, OT could not account for **opacity**, i.e. the situation in which an apparent phonological rule is contradicted by the output data and yet deemed valid because the contradicting data are the result of a later rule. Earlier critics of SPE had rejected rule ordering as a valid mechanism and thus concluded that most opaque rules are actually morpho-lexical rules. Proponents of OT, however, remained faithful to SPE and thus proposed numerous ways to have their cake and eat it (reject ordering and yet account for opacity). In recent influential versions, extrinsic ordering (of blocks of constraints) has been brought back, so that OT no longer can do with just two levels (see Kiparsky's 2000 Stratal OT and the model proposed in McCarthy 2007).

5.5.3   *Back to neutralization*

We now need to make explicit that allophonic rules that are neutralizing can in principle be analyzed at the phonemic level as well, precisely because neutralizing rules output phonemes (or, as in the case of flapping can be stated as outputting phonemes which then undergo an allophonic rule). If we require that phonetic rules have no exceptions we could not relegate morpho-lexical rules or phonotactically-motivated phonemic rules that have exceptions to the phonetic level, but, unless we require that phonemic rules must have exceptions, we cannot ban automatic rules from the phonemic level, precisely because they can be analyzed as producing phonemes as outputs.

Turning back to final devoicing in Dutch one thus can explore two possible analyses, which I will discuss in detail here. In section 5.4., I assumed that final devoicing (FD) follows from a phonetic constraint.

(27) * [-son, +voice])$_{\sigma op}$

To analyze FD as an allophonic rule I will first eliminate an analysis that does *not* work and then present the correct analysis which I will compare to a possible phonemic analysis. We could assume that alternating obstruents are unspecified for voice and postulate the value adding rule in (28b):

(28) a. /hɔnt/ 'dog' SG   /hɔndən/ 'dog' PL
          [∅voice]   [∅voice]

    b. [∅voice] → [+voice] /$_\sigma$(_
        [–voice] elsewhere (=default)

In this analysis FD is actually analyzed in terms of a voicing rule. The problem is that words that always have a voiceless obstruent (see (29)), while being predictably voiceless when the obstruent is syllable final, cannot be unspecified for [voice] because that would trigger the rule in (28b).

(29) [kɑt] 'cat' SG   [kɑtən]   'cat' PL
     [wɑnt] 'mitten' SG   [wɑntən] 'mitten' PL

We must thus specify non-alternating obstruents as [-voice], in a sense, treating them as exceptions to the voice-alternation. This undesirable consequence suggests that the real generalization is about final devoicing after all and not about 'medial voicing'. This then must lead to a different analysis in which we specify alternating obstruents as lexically [+voice]. When occurring without an affix such forms violate the phonetic constraint in (27) which triggers a repair rule that changes [+voice] to [-voice]. Since non-alternating voiceless stops could be specified as [∅voice], we could then say that the final devoicing rule can fill in [-voice], while this same rule is allowed to change [+voice] to [-voice] in a 'derived environment' (where 'derived' here means that the obstruent in question is followed by a word boundary.) This requires the following formulation (i.e. one that does not require either [∅voice] or [+voice] in the input):

(30) [-son] → [-voice] / _)$_\sigma$

I stress that if final devoicing is seen as serving a phonetic constraint, there is no phonemic illformedness in ending in a voiced obstruent.

However, since final devoicing is neutralizing, we could also analyze the alternation in terms of a phonemic repair rule, in which case we regard constraint (28) to refer to the phonemic level. Everything else can be the same as in the phonetic analysis.

Apparently we have to accept that rules that output phonemes and that are not dependent on morphological information *can* be analyzed in two ways, i.e. that they are ambiguous. This ambiguity does not have to be regarded as a theoretical flaw. Rather it provides a channel for neutralizing allophonic rules to become phonemic rules and then, subsequently, become tolerant to exceptions.

However, in this case, it is unlikely that final devoicing case is a phonemic rule. Recently, it has been argued in a variety of cases that alleged neutralizing rules are, in fact, not neutralizing at all. For example, it has been argued for various languages that have final devoicing that de devoicing is not complete. Whether alleged neutralizing rules are truly ambiguous or not must be decided on a case by case basis.

6. Concluding remarks

In this chapter I have been unable to provide a complete overview of phonology and several important subjects have not been touched upon. In the domain of allophonic rules, I have not considered the issue of **stylistic rules**, also called fast speech rules, casual speech rules or allegro rules. We could also include in this class, the variable rules proposed by Labov (1966). Where in the phonological model do we account for such differences? There are two ways to go. If we assume that all **registers** share phonological properties up to the phonetic level, we could assume an additional 'stylistic' module which produces a set of stylistic phonetic representations, based on a set of variables that determine these registers, which then all connect to the implementation module, or we could build these variables directly into the phonetic implementation module? The fact that word forms are subject to stylistic rules raises the further question which pronunciation is taken to be the basis for the lexical representation. I refer to Linell (1979) for arguments that the so-called **careful pronunciation** serves as the basis for lexical representations, which, in fact, is an implicit practice observed by almost every linguist.

This leads us to further questions regarding the nature of lexical entries? I have presented a model in which lexical representations are phonemic representations, i.e. they abstract away from allophonic properties. The issue remains whether such representation are impoverished (in that redundant specification are left out). A further question is whether lexical

representations are *word forms* or can be more abstract, including morphemes such as /hɔnd/. If, as Linell (1979) argues, lexical representations must be word forms, the lexical form of words that alternate in final voicing cannot be the form ending in a voiced obstruents since this form never occurs as an actual word form which enforces an representation such as /hɔn{t/d}/ (cf. 17), where 't/d' means that the 't' alternates with a 'd' in related word forms. This strikes me as undesirable.

It was stated in section 3 that all phonological theories deal with derivations and representations. In this chapter I have focused on derivational issues. I have assumed that representations are discretely compositional with features organized in segmental units at the phonemic and phonetic level. I have mentioned constraints that govern wellformed combinations of such segments (at both levels) that refer to hierarchical units such as syllables and feet. I refer to Ewen and van der Hulst (2000) for an overview of representational aspect up to the word level and to Nespor and Vogel (1986) for higher level organization.

References

Anderson, John M. & Colin J. Ewen (1987) *Principles of Dependency Phonology*. Cambridge: Cambridge University Press.

Anderson, Stephen R. (1975) On the interaction of phonological rules of various types. *Journal of Linguistics* 11: 39-62.

Anderson, Stephen R. (1985) *Phonology in the twentieth century: Theories of rules and theories of representations*. Chicago: University of Chicago Press.

Anderson, Stephen and David Lightfoot (2002) *The language organ. Linguistics as cognitive physiology*. Cambridge: Cambridge University Press.

Archangeli, Diana (1984) *Underspecification in Yawelmani phonology and morphology*. Ph.D. dissertation, MIT. Published 1988, New York: Garland.

Bell, Alexander M. (1867) *Visible speech: The science of universal alphabetics*. London: Simplin, Marshall and Co.

Bybee, Joan (2001) *Phonology and language use*. Cambridge: Cambridge University Press.

Chomsky, Noam and Morris Halle (1965) Some controversial questions in phonological theory. *Journal of Linguistics* 1: 97-138.

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

Dresher, B. Elan (2009) *The contrastive hierarchy in phonology*. Cambridge: Cambridge University Press.

Ewen, C. and H. van der Hulst (2000) *The phonological structure of words. An introduction*. Cambridge: Cambridge University Press.

Goldsmith, John (1993) Harmonic phonology. In John Goldsmith (ed.), *The last phonological rule*, pp. 21-60. Chicago and London: The University of Chicago Press.

Halle, Morris (1959) *The sound pattern of Russian: A linguistic and acoustical investigation*. The Hague: Mouton.

Harris, John (1987) Non-structure preserving rules in lexical phonology. *Lingua* 72.4, 255-292.

Hjemslev, Louis (1953) *Prolegomena to a theory of language*. Translated by Francis Whilfield. Madison: University of Wisconsin Press. [Original 1943]

Hockett, Charles (1955) *A manual of phonology*. Bloomington: Indiana University Press.

Hooper, Joan (1986) *An introduction to Natural Generative Phonology*. New York: Academic Press.

Hudson, Grover (1974) The representation of non-productive alternations. In John Anderson and Charles Jones (eds.), *The First International Conference on Historical Linguistics* [Volume II], 203-229. Amsterdam: North Holland.

Hulst, Harry van der (2013) The discoverers of the phoneme. In Keith Allan (ed.) *Oxford Handbook of the History of Linguistics,* pp. 167-191. Oxford: Oxford University Press.

Itô, Junko (1986) *Syllable Theory in Prosodic Phonology*. Doctoral dissertation, University of Massachusetts, Amherst. [Published, 1988, Outstanding Dissertations in Linguistics series. New York: Garland.]

Jakobson, Roman, C.Gunnar M. Fant and Morris. Halle (1952) *Preliminaries to speech analysis – The distinctive features and their correlates*. Second Edition. Cambridge, MA: MIT Press.

Kaye, Jonathan, Jean Lowenstamm and Jean-Roger Vergnaud (1990) Constituent structure and government in phonology. *Phonology* 7: 193-232.

Kiparsky, Paul (1982) From cyclic phonology to lexical phonology. In Harry van der Hulst and Norval Smith. eds. *The structure of phonological representations*, pp. 131-177. Dordrecht: Foris.

Kiparsky, Paul (2000) Opacity and cyclicity. *The Linguistic Review* 17, pp. 351-367.

Kisseberth, Charles (1970) On the functional unity of phonological rules. *Linguistic Inquiry* 1: 291-306.

Koskenniemi. Kimmo M. (1983) Two-level morphology: a general computational model for word-form recognition and production. Doctoral dissertation, University of Helsinki.

Labov, William (1966) *The social stratification of English in New York City.* Washington, DC: Center for Applied Linguistics.

Lahiri, Aditi and Reetz, Henning (2010) Distinctive Features: Phonological underspecification in representation and processing. *Journal of Phonetics* 38: 44-59.

Lamb, Sidney (1966) Prolegomena to a theory of phonology. *Language* 42: 536-573.

Levelt, Willem J. M. (1989) *Speaking. From intention to articulation.* Cambridge MA: MIT Press.

Liberman, Mark and Janet Pierrehumbert (1984) Intonational invariance under changes in pitch range and length In Mark Aronoff and Richard T. Oehrle (eds) *Language and sound structure*, pp. 157-233. Cambridge, Mass.: MIT Press.

Lillo-Martin Diane and Wendy Sandler (2006) *Sign language and linguistic universals*. Cambridge: Cambridge University press.

Linell, Per (1979) *Psychological reality in phonology*. Cambridge: Cambridge University Press.

McCarthy, John J. (2001) *A Thematic guide to Optimality Theory*. Cambridge: Cambridge University Press.

McCarthy, John J. (2007) *Hidden Generalizations: Phonological Opacity in Optimality Theory*. London: Equinox.

Nespor, Marina and Irene Vogel (1986) *Prosodic phonology*. Dordrecht: Foris publications.

Pierrehumbert, Janet (1990) Phonological and phonetic representation. *Journal of phonetics* 18: 375-394.

Pierrehumbert, Janet (2001) Exemplar dynamics: Word frequency, lenition and contrast. In Joan Bybee and Paul Hopper (eds) *Frequency and the emergence of linguistic structure*, pp. 137–157. Amsterdam and Philadelphia: John Benjamins.

Pike, Kenneth L. (1947) *Phonemics: A technique for reducing languages to writing*. Ann Arbor: University of Michigan.

Prince, Alan and Paul Smolensky (1993) Optimality theory: Constraint interaction in generative grammar. *Technical report #2 of the Rutgers center for cognitive science.* Piscataway, New Jersey: Rutgers University. [Published in 2004. Malden, MA and Oxford: Blackwell.]

Pullum, Geoffrey and William Ladusaw (1986) *Phonetic symbol guide*. University of Chicago Press.

Rogers, Henry (2005) *Writing systems. A linguistic approach*. Oxford: Blackwell Publishing.

Scobbie, James (1997) *Autosegmental Representation in a Declarative Constraint-based Framework*. New York and London: Garland press. [Original dissertation: Attribute Value Phonology. Edinburgh, 1991]

Strauss, Steven (1982) *Lexicalist phonology of English and German*. Dordrecht: Foris Publications.

Taylor, John (2003) *Linguistic categorization*. Oxford: Oxford University Press.

Trubetzkoy, Nikolai S. (1939/1969) *Principles of phonology*. Translated by Christiane A. M. Baltaxe. Berkeley and Los Angeles: University of California Press. Originally published 1939 as *Grundzüge der Phonologie*. Göttingen: van der Hoeck and Ruprecht.