Underlying Representations

The content of underlying representations. A basic issue regarding underlying forms is: what are they made of? We have so far treated them as segments represented as letters. Chapters 6 and 11 revisit representations, showing that segments are not atomic units but are defined in terms of a small set of phonetically defined universal features. Correspondingly, underlying forms are composed of sets of features, and not unanalyzable segments per se.

Chapter 10 addresses the famous abstractness issue, concerning the extent to which underlying forms can differ from surface forms. We have partially addressed that issue here, showing that the underlying form is not necessarily a substring of any one word that contains the morpheme. Hence in Palauan, the two vowels of /daŋob/ are never both present as full vowels in any single instantiation of the root; in Tonkawa, the root /picena/ is never directly seen in that form, since one of the vowels always deletes. The main question addressed in Chapter 10 is whether an underlying form may even contain segments which are never pronounced as such in the language.

A further issue in phonological theory has been the question whether underlying representations must be "in principle pronounceable". For instance, an underlying form of the type /sa5t%o¶£/ would be typically precluded as an underlying form, since the symbols [5 % ¶ £] cannot be pronounced, so it would be natural to assume that such elements cannot be part of a phonological representation. But it is often — though not universally — assumed that underlying representations can have so-called "archiphonemes" such as /B/ which conventionally represent a segment that is not definitely voiced as in /b/, and is also not definitely voiceless as in /p/. As such, an archiphoneme is also unpronounceable, though in a different way from supposed /¶/, in that it is *almost* pronounceable, and in practice always becomes pronounceable by making a definite decision, by applying some rule, as to whether the segment is [p] or [b]. The difference between /¶/ and /B/ is that the former has *no* pronounceable attributes.

Morphology and morphophonemics. We have seen that in the taxonomic tradition, phonological rules of the type discussed in this chapter were referred to as morphophonemic rules (a name still used occasionally to describe phonological processes). The implication is that there is a connection between morphology and such phonological processes. Allophonic rules do not neutralize distinctions between underlying segments, so the underlying form can be automatically inferred from the surface form — per the condition of biuniqueness. But with the output of a neutralizing rule (such as those presented in this chapter) it is impossible to know which of two (or more) segments a surface segment X derives from, if you only look at the word itself, even if you know the system of rules: you can't know in Russian whether [porok] comes from /porog/ or /porok/. You must look at a different instantiation of that very segment, but appearing in some other context, in order to determine what the underlying segment is. Most often, the other context which provides the crucial evidence must be a word with a morphological relation to the word in question, such as an infinitive of a verb, compared to a 1sg present tense form (as was needed to distinguish /h/ and /p/ in Kikerewe, which neutralize to [p] after a nasal). For that reason, neutralizing phonological alternations are most often associated with different morpho-

logical forms of words: hence the term "morphophonemics", which is the selection of a particular (taxonomic) phoneme in some context.

In the taxonomic approach, morphophonemic rules are said to select **allomorphs** of a morpheme. Analogous to the phoneme being seen as the set of all phones that realise the phoneme, the morpheme is seen as being made up of all of the surface (phonemic) variants of the morpheme, thus the plural morpheme in English would be composed of {/s/, /z/, /ɨz/} (there could be other things included in the morpheme, such as /n̩/ as in "oxen"), and the plural morpheme ||S|| would have the allomorphs /s/, /z/, and /ɨz/. There could also be auxiliary symbols used to represent processes. As an example of the latter kind of symbol, the morphophoneme ||≃|| is used in Miller's (1965) grammar of Acoma to indicate that the preceding vowel is deleted; ||M|| indicates a labial nasal which is retained before a particular list of suffixes and is deleted elsewhere; ||B|| symbolizes a consonant that is realised as [b] before certain vowel-initial suffixes, and deletes otherwise. In the case of Tonkawa, the morpheme which might be symbolized as ||notoxo|| "hoe" is realised by the set of allomorphs {/ntox/, /ntoxo/, /notx/, /notxo/}, and the morpheme ||picena|| represents the set {/pcen/, /pcena/, /picn/, /picna/}. Morphophonemics would refer to some set of rules that select particular variants in some context, just as there are rules that select the variants of the English verb morpheme ||be||, namely {/ɪz/, /wəz/, /wr/, /æm/, /ar/}. The problem with the taxonomic view of "morphophonemics" and "allomorphs" is that it fails to distinguish between variations in form that are due to a phonological rule such as vowel-harmonic suffix variants in Finnish or phonetically-driven changes in the plural suffix in English, versus suppletive alterations as in the verb "be".

Not all neutralizing phonological rules involve changes induced by varying the morphological context: the defining context for a rule might be surrounding words in a sentence. We saw this in Korean, with a rule nasalizing stops before a nasal consonant. This rule applies even at the sentence level, so the pronunciation of the word /pap/ "rice" varies between [pap] and [pam], depending on which word follows: the morphological form of the word /pap/ — uninflected noun — remains constant in [pam məkəra] "eat rice" and [pap twemnita] "served rice", even though the pronunciation of the word changes. There is a rich literature addressing such neutralizing phonological rules at the sentence level: see for example Kisseberth & Abasheikh (1974), Clements (1978), Odden (1987). Most examples of sentence-level phonology involve just the consideration of a word an an immediately preceding or following word; however, Odden & Robert-Kohno (1999) and Odden (2000) discuss phrasal processes in Kikamba and Zinza which apply across unbounded sequences of words. See Inkelas & Zec (1995) and references therein for various theoretical perspectives on phrasal phonological rules.

The relation of phonology to morphology. A different view of the relationship between phonology and morphology is presented in the theory of Lexical Phonology (Kiparsky 1982, Kaisse & Shaw 1985, Mohanan 1986). In the theory presented in this book, it is assumed that the underlying form contains *all* of the morphemes present in a word, which are present at once, and they have whatever shape they are assigned in their dictionary entries. In the Lexical Phonology model, phonology and morphology are separate but bidirectionally interacting components in a grammar. Rules are assigned to particular strata or levels in both phonology and morphology, which allows words to be built up in stages so that *some* phonological rules may apply early in a

derivation, and then may be followed by morphological operations of affixation. Certain morphological operations are defined for the first stratum of the morphological component — Level 1 — so in English, affixation of the derivational affixes *-ion* and *-ity* takes place at Level 1, and other operations such as affixation of *-ness* and *-ing* are defined at Level 2. Similarly, certain phonological rules are defined for Level 1 but not Level 2, for example stress assignment and Trisyllabic Shortening (which accounts for the vowel alternation in *profane ~ profanity*) which apply only when Level 1 affixes are added.

Under this model, the output of Level 1 morphology undergoes the rules of Level 1 phonology. The result derived from Level 1 phonology is then resubmitted to the morphological component, and at Level 2 other affixes are added, with the resulting form again being run through the phonology. Thus the derivation of a word, in the view of Lexical Phonology, involves a shuttling back and forth between the morphological component and the phonolological component. Accordingly, each input to the phonology (at Level 1 vs. Level 2) would be the underlying form, at that level. Further consideration of Lexical Phonology will be given in Chapter 5, in connection with rule ordering and the cycle.

Morpheme structure conditions and rules. One of the more resilient problems regarding underlying forms is that they seem to exhibit regular properties. For instance, there are no morphemes in English which begin with sequences such as **bn*, **km*, **gn*; in Korean, no noun stems end with the consonants *p'*, *t'*, *k'*, or, for that matter, *t*; Russian has no underlying consonants plain /č/, palatalized /ž^{y}/, or /γ/. If all regularities about languages are to be stated in a grammar (a major assumption worth critical thought), how are these facts to be expressed?

One answer, exemplified by Halle 1959, is that there are **morpheme structure rules** (MSR) in a grammar which perform specific repairs on any string not conforming to such rules. Thus there could be a rule of Russian that would converts any non-palatalized (phonetically non-existent) /č/ into palatalised [č^{y}], whioch would account for the fact that there are no non-palatalized alveopalatal affricates in Russian. Paired with the assumption that there are such rules is that certain properties would be left blank underlyingly (resulting in archiphonemic representations such as /C̆/ with a segment which is neither explicitly palatal nor explicitly non-palatal), so in Russian the palatalization feature would not be underlyingly specified for the alveopalatal affricate, and would be filled in by a rule. An alternative to such a rule is **morpheme structure conditions** (MSC) as proposed by Stanley 1967 which do not 'actively' change underlying forms, but rather 'passively' impose requirements on what is a well-formed morpheme. In the MSC approach, underlying representations will contain fully specified segments, thus the voiceless alveopalatal affricate of Russian is definitively palatalized, and is not left in a "blank" state for palatalization. MSCs do not state what would happen to any segment contradicting the requirements of the condition, but typically violations are considered to impose a "cost" on the grammar (see the commentary on Chapter 6, 8 for discussion of "cost").

The major criticism of the MSC approach is to so-called "duplication problem" identified in Kenstowicz & Kisseberth 1977, that MSC's are often the same as active segment-changing rules motivated by the existence of paradigmatic alternations. For example in Kiruundi, all vowels within roots are underlyingly long when followed by a sequence of a nasal plus a consonant — an example is [riind-a] "watch", which can be explained by a MSC requiring vowels to be

long before NC. There is also an active alternation in vowel length triggered by combination of underlying short vowels plus NC sequences, which explains the variation in the length of the infinitive prefix /ku-/ seen between [ku-ror-a] "to look at" vs. [kuu-n-dor-a] "to look at me". The MSC on underlying length duplicates the phonological rule which must exist anyway.

It does not solve the problem to derive morpheme-internal vowel length in [riind-a] by simply applying this rule to an underlying short vowel, assuming that all vowels are underlyingly short before tautomorphemic NC. Under the principle that all regularities must be stated in the grammar, this just replaces the duplication problem with a "mirror-image duplication" problem, that if there is a rule deriving X in a context, there must be a MSC ruling out X in that context. An alternative would be to say that the assignment of vowel length within morphemes before NC is random, so that the underlying form of [riind-a] might be /rind-a/ but the underlying forms or [baamb-a] might be /baamb-a/. No facts of the language give evidence for supposing /rind-a/ vs. /baamb-a/, so the question of how speakers could learn the specific underlying forms or roots is unanswered.

One contemporary trend in phonology, Optimality Theory, theoretically precludes the possibility of a grammar with statements about underlying forms. This has resulted in the principle "Richness of the Base" (Prince & Smolensky 1993), variously understood as saying "any imagineable input must map to some output" or "any string must map to a well-formed output", ideas which have in common the denial that a grammar states regularities about underlying forms. One significant shift of emphasis in that approach, discussed under the rubric "Stampean occultation" and "lexicon optimization" (see also Inkelas 1994), is that some observable regularities may not be part of a formal grammar. The lack of any underlying /ü/ in the lexicon of English words might thus not be accounted for in the formal theory of grammar, but emerges as a byproduct of language acquisition — in a nutshell, there is no reason to assign /ü/ to the underlying form of any words of English. This seems to be a major break with previous tradition, since it rejects the idea that all generalizations about a language need to be explicitly encoded in a grammar. Another way to view ROTB is that it entails that a grammar must contain explicit devices (constraint rankings) which prohibit "impossible" outputs such as *[stiyŋk] in English, no matter what the underlying form is. This question remains unresolved at present.

Readings

Clements, G. N., 1978, Tone and syntax in Ewe. in D. J. Napoli (ed.) *Elements of tone, stress, and intonation*, 21-99. Georgetown University Press: Washington)

Halle, M. 1959. *The sound pattern of Russian*.

Harris, Z. 1942. Morpheme alternants in linguistic analysis. *Language* 18:169-180.

Hockett, C. 1947. Problems of morphemic analysis. *Language* 23: 321-343.

Inkelas, S. & D. Zec. 1995. Syntax-phonology interface. J. Goldsmith (ed.) *Handbook of Phonological Theory*, 535-549. Oxford: Blackwells.

Kaisse, E. & P. Shaw. 1985. On the theory of lexical phonology. *Phonology* 2: 1-30.

Kenstowicz, M. & C. Kisseberth. 1977. *Topics in phonological theory*. Academic Press: New York.

Kiparsky, P. 1982. Lexical morphology and phonology. I.-S. Yang (ed) *Linguistics in the morning calm*, 3-91. Seoul: Hanshin.

Kisseberth, C. & M. Abasheikh. 1974. Vowel length in Chi-Mwi:ni — a case study of the role of grammar in phonology. A. Bruck, R. Fox & M. LaGaly (eds.), *Papers from the parasession on natural phonology*, 193-209. CLS.

Mohanan, K. P. 1986. *The theory of Lexical Phonology*. Reidel: Dordrecht.

Odden, D. 1987. Kimatuumbi phrasal phonology. *Phonology* 4, 13-36

Odden, D. & R. Roberts-Kohno. 1999. Constraints on Superlow tone in Kikamba. R. Kager & W. Zonneveld (eds), *Prosodic and phrasal phonology*, 145-170. Nijmegen University Press, Nijmegen (w. R. Roberts-Kohno).

Odden, D. 2000. The phrasal tonology of Zinza. *Journal of African languages and linguistics* 21: 45-75.

Prince, A. & P. Smolensky. 1993. Optimality theory: constrarint interaction in generative grammar. Report no. RuCCS-TR-2. Rutgers University: New Brunswick.

Stanley, R. 1967. Redundancy rules in phonology. *Language* 43: 393-436.