

WHERE IS THE NATURAL PHONOLOGY PHONEME IN 2007?

Geoffrey S. Nathan

Wayne State University, USA
geoffnathan@wayne.edu

ABSTRACT

This paper reviews the status of the phoneme, the basic linguistic unit of phonology as it is understood within Natural Phonology, and the ways in which it contrasts with current mainstream theories—particularly most versions of Generative Phonology on the one hand, and various usage-based view on the other. It concludes by pointing out a number of aspects of phonological acquisition, storage and processing which indicate that there is still a role for the traditional Baudouin/Sapir/Stamppe-Donagan view of fully-specified acoustic/articulatory idealized targets.

1. INTRODUCTION

The contributions to this workshop have dealt with a number of distinct issues involving how Natural Phonology (NP) relates to other theories prevalent in the early part of the twenty-first century. With the exception of a brief, and elegantly expressed discussion in Dziubalska-Kołaczyk [1] the central concept of the phoneme itself is not discussed by any of the contributors. This contribution will discuss the status of this old, but battered idea in the context both of its status within ('outsider') Generative Linguistics but also within ('insider') versions of functionalist linguistics that have become quite popular in the non-generative arena over the past ten to fifteen years.

As Dziubalska-Kołaczyk points out, NP's view of the phoneme is quite different from the traditional structural *and* generative view. Differing from the structuralist view, phonemes are not defined by their opposition within a system, but are rather defined *psychologically*, as the mode of perception, storage and production of speech by native speakers of a language.

2. PHONEMES ARE NOT CONTRASTS

Concepts such as 'contrast' can be used as discovery procedures—as a way for a linguist to get 'into' the system if the investigator is not a native speaker, but contrast and complementary distribution do not define phonemes. This view, the rejection of 'discovery procedures' as a *definition* of the phoneme, is not unique to NP, of course, but formed the basis of Chomsky and Halle's rejection of structuralist phonology in the 60's. Because phonemes are not defined by contrast, therefore, concepts such as 'archiphonemes'—sounds that do not 'know' what they are (i.e. are indeterminate between otherwise contrasting phonemes) have no place in NP. Since phonemes are how speakers perceive sounds, the concept of a different kind of sound that is not quite a sound is incoherent with the NP view of what phonology is all about.

3. PHONEMES ARE NOT FEATURES

However, because phonemes are *sounds as perceived*, this means that they are auditory/motor *images* of sounds per se. not abstract specifications for sounds. Thus, contrary to what is usually believed in most (but not all) generative phonologies, phonemes are not 'merely' lists of features. And particularly, they are not underspecified lists of features. It is important to how NP works that phonemes are real (although mental) sounds, fully specified. What makes them phonemes, rather than just records of how speakers actually speak, is the existence of processes. Phonemes are very rarely pronounced as stored, but instead are modified either to fit their environments (lenitions) or in contrast to their environments (fortitions). Phonemes are specific mental targets, so that we could imagine pronouncing a phoneme such as /i/, whereas a specification such as [V, +hi, -back] is not pronounceable, and hence, not what NP means by a phoneme, because the feature list

lacks specifications for nasality, tenseness, length and so on. A Natural Phonology phoneme is fully specified for all possible features—it's just that when it is actually pronounced in context it doesn't necessarily come out that way.

It is this particularly intricate, and detailed view of how speech perception and production works that determines NP's view of the phoneme. Stampe argued that perception is not mere recording of external events, but rather, at least for the perception of human action, is a perception of the *intention* of the other. As Dziubalska-Kolaczyk [1] says

a phoneme...is an underlying *intention* shared by the speaker and the listener (who are always "two in one"). The shared knowledge of intentions guarantees communication between the speaker and the listener within a given language, even if the actually pronounced forms diverge substantially from what is intended...In other words, phonemes are fully specified, pronounceable percepts.

What this means is that speakers automatically (and generally unconsciously) adjust their intentions to make them fit their environments (generating allophones 'on the fly') and equally unconsciously 'undo' the processes as they hear their fellow speakers produce sounds that they themselves would not have intended. So, for example, native speakers of American English, hearing a nasalized, diphthongal [ɛ̃] assume the existence of a following /n/, even if it is not pronounced (as in 'paint', for example). Similarly, Americans, on hearing [mæ] hear a /t/ between the two vowels in 'matter', since they know they 'delete' flaps intervocally in casual speech. They know that if they had intended a /t/ it would have disappeared, so when someone produces a word that they recognize as having an intended /t/ they hear the /t/, whether pronounced or not (and, if you ask someone 'What did he say?', they would answer [mæt] without noticing the difference.

4. THE NEW FUNCTIONALIST CHALLENGE

In the past ten years or so a new alternative view of phonology, dealing with both acquisition and storage, has arisen, both within the larger functionalist theory known as Cognitive Grammar and among a group of phonologically-interested phoneticians.

Cognitive Grammar (CG) is a functionalist theory of language that rejects many of the fundamental tenets of Generative Grammar. It does not hold that language is determined by autonomous mental structures, but rather that all language processing (acquisition, storage, perception and production) follows the same principles as other aspects of human mental behavior. Since its introduction in the mid eighties it has evolved so that some early foci have become less important and other operating principles have come to the fore instead.

The early version of CG emphasized the ability of humans to categorize their world, and particularly was interested in the nature of that categorization. Drawing on the work of Eleanor Rosch and her colleagues, it assumes that linguistic categories (like all other categories) have a radial prototype structure (see Lakoff [9] for the classic early description of some linguistic categories at various levels). Langacker, in his early work ([10]) argued that linguistic units formed a continuum and that hence there was no clear-cut division between morphology and syntax, nor between syntax and lexicon. Lexical units could span multiple words, for example (such as in the case of idioms). In addition, he rejected the fundamental dichotomy between grammar and lexicon, arguing that totally productive constructions are stored as wholes, *in addition to being generated by rules*. Thus he argued that frequently produced English noun plurals were stored as such, even if they involved totally predictable morphology.

Langacker [10], [11] also introduced the Content Constraint, which argued that there were no meaningless units in language (such as pleonastic 'it' in extraposition sentences) and that paraphrases always had at least minimally different meanings which were a consequence of the meaning of the different constructions that were contained therein. Thus he showed how the subject-raising construction ('John seems to be happy') means something different from the extraposition construction ('It seems that John is

happy') by virtue of the different words the sentences contain (e.g. presence vs. absence of 'it') and the fact that 'John' is subject of the main clause in one sentence and subject of an embedded clause in another. Some of the differences are related to information structure and 'scene-setting', and thus perhaps more pragmatics than semantics, but CG in general does not recognize a strict dichotomy between those two realms either.

Within the theory of CG early work on phonology by this author attempted to place the NP phoneme into the view of categorization espoused by Lakoff (Nathan [12],[13]).

Later developments in CG concentrated more on how language was acquired. Beginning with the Content Constraint, Langacker and others ([2]) argued that language was acquired solely through exposure to linguistic input and general cognitive principles of categorization and abstraction. Consequently there was no need for independent linguistic principles. Since Langacker argued that lexical items were stored with all morphology intact, and might or might not even have the morphology analyzed, at least at first, language acquisition was simply a case of storing large numbers of individual instances, then extracting generalizations from the stored instances, while not discarding the instances stored. Thus, we might hear *dogs*, store it as heard (i.e. as a single, plural word), and perhaps later generate a separate stored item 'dog'. Or perhaps not.

This view, which came to be called a 'usage-based' theory (Barlow and Kemmer [2]), spread naturally to phonology as well. Bybee and others ([3], [4], [5]) have argued that words (and probably larger units too) are stored exactly as heard (and as produced). This would indicate massive storage needs, but there is evidence that the brain is indeed able to store enormous numbers of incidences. Thus when children acquire their language they simply store everything they hear, in great phonetic detail. Langacker, Kemmer, Bybee and others have argued that abstract schemas can be extracted from these stored units, and that phonemes, syllables and other higher order units are simply generalizations over stored individual instances.

This highly concrete view of phonological processing is also proposed by a group of

researchers with phonetic backgrounds. Pierrehumbert [15] have been developing a similar view of phonology, using the recently developed concept of categorization that has come to be known as Exemplar Theory (ET). Within this theory, large numbers of individual instances are all stored, but that similar instances are stored, in some sense, 'on top of' each other, somewhat analogous to tracing a path through a park. As more and more instances in the same place are stored a visible path emerges where the grass is worn away, but all the other paths that do not coincide exactly with the bare earth are still there also, and available to be perceived and produced. However, the more frequent instances are 'stronger' and hence more likely to be produced, and instances that are heard and are similar to the 'beaten path' are likely to be assimilated to it.

5. A NATURAL PHONOLOGY RESPONSE

Thus, the crucial way in which these Usage-Based theories differ from NP's view of phonological processing is that words are stored exactly as perceived (and motor memories of differing productions are also stored individually, with all their variations). NP, on the other hand, argues that words are perceived as intentions, and instantaneously 'respelled' with phoneme targets as components.

There are numerous arguments that could be developed to sway linguists in one direction or another in this controversy. As evidence for the Usage-Based view there is the fact that people seem to be able to recognize new words faster if they had heard them previously spoken by the same voice as the test version. And we certainly can recognize when someone says *dog* 'just like Uncle Joe', which indicates that we do indeed store speaker-specific information 'within' the lexical entry.

On the other hand, there is still a great deal of evidence that lexical entries are indeed stored as a list of phonemes in the NP sense.

First, there is production evidence. Speech errors do not scramble allophones but phonemes. This fact is so well known that no examples are needed. If words were indeed stored phonetically, displaced segments would come out 'wrong' for their environment and 'Scotch tape' would be

pronounced [st ɔ̃ ke p] with an aspirated [t] but a velar (rather than palatal) and *unaspirated* [k]. As Donegan and Stampe noted thirty years ago, this never happens.

Second, there is the fact that learning to spell with an alphabetic system is much easier than learning to spell with a word-based system such as Chinese/Japanese kanji. Further, children acquiring spelling in languages that use a relatively straightforward phoneme-grapheme correspondance pattern learn earlier than those with more abstract systems such as English (see Treiman [20] for extensive discussion). But even more important is the fact that children have no difficulty perceiving the ‘t’ sound in *stop* as the same as the one in *top*. Similarly they don’t protest when the same letter is used for the ‘l’ sound in *leaf* and in *feel*, even though at least some speakers pronounce the syllable-final /l/ as more vowel than consonant (the present writer being one such person). Not only do we never hear of children having such problems, but the majority of reading teachers are unlikely to know that the two sounds are different, since elementary phonetics and the concepts of phonemes and allophones are rarely required of elementary school teachers. And certainly beginning linguistics students have some difficulty hearing allophones, and often do narrow transcription by ‘rule’ rather than actually hearing the surface forms.

Secondly, I have argued elsewhere ([14]) that sociolinguistic variation in sounds in synchronic sound shifts (such as the New York vowel shift studied by Labov) and the Northern Cities vowel shift studied by Labov argue that phonemes are uniform articulatory/acoustic units, and that any word containing the relevant sound contains the same set of articulatory/acoustic specifications. We know that these heavily studied sound changes apply uniformly to all words containing the relevant sound, or do not apply, depending on social and frequency factors for specific words, but for any given phoneme either the *same* sound change applies, or doesn’t. There are not different sound changes for the same phoneme for different words--we do not find the /æ/ in *bad* backing while the /æ/ in *back* raises. Clearly there is one /æ/, and some words participate in its changing while others don’t.

Finally, there is the fact that children acquiring language are much more active participants in the sound-learning process than the Usage-Based theory would suggest. Examples of child innovations are, of course, legion, and the classic work of Jakobson [8] pointed out the similarities between child substitutions and historical changes. A child that some colleagues and I have been studying, who is learning American English and Spanish simultaneously, went through a period where he systematically replaced all Spanish /r/’s, whether tapped or trilled (recall that these contrast in Spanish, at least intervocalically) with /l/’s. Neither his mother nor his father speak this way (one speaks Asturian Spanish, the other Northeast American English). There is no ‘usage’ that could serve as the source of this well-known and frequent sound change.

6. Conclusions

How does NP stand among competing theories in 2007? While it has very few adherents in North America, it is still alive in Europe. How does it compare with the major linguistic theories competing for adherents? Unlike Generative Grammar it is not a formal theory, and does not make the claim that without formalization the theory is not testable. Similar to at least one version of Generative Grammar, Optimality Theory, it understands that the synchronic state of any language is a result of a dynamic tension between competing cognitive demands—a need for physical and cognitive efficiency and a need for auditory and cognitive clarity. Efficiency is essentially the speaker’s goal for him/herself, while clarity is the speaker’s goal for the listener. NP’s view of the phoneme is that the storage of words using a finite set of phonological units (phonemes) represents language’s ‘solution’ to this dilemma. Each phoneme has been shaped to be maximally perceptible and the inventory has been shaped to be relatively small (rarely fewer than fifteen, rarely more than forty)—enough to express the entire lexicon, but not too many to master or keep track of. Thus NP is an explanatory theory (as Donegan and Stampe argued in their seminal paper ([6])). In this respect it differs from Generative Grammar in that the principles of GG are considered sui

generis, and explicitly not derivable from other cognitive or social principles.

On the other hand, NP differs from at least some current versions of CG (the leading functionalist theory at the moment) in that it ascribes an active role for the speaker/hearer in the perception and production of speech. By active I mean that the listener hears speech (of whatever sort--adult, child, different dialect, different language, even birdsong or ambient noise) and (metaphorically speaking) thinks 'what sounds would I have made that would come out sounding like that?' Thus all speech is encoded as a sequence of native language phonemes, whether just talking to a friend or attempting to communicate the sound of a bird species or ambulance siren (French police cars, for example are described as saying 'honk bee, honk bee')

What distinguishes NP's view of phonemes from the Usage-based theory is that principles of symmetry, efficiency and clarity shape the basic units as the language is acquired. The units are developed not merely as an abstraction of surface forms, but as an active construction of intentions from the surface forms as they 'come in' to the organism.

7. References

- [1] Archangeli, D. (1988). Aspects of underspecification theory. *Phonology*, 5(2), 183–208.
- [2] Barlow, M., & Kemmer, S. (Eds.). (2000). *Usage-based models of language*. Cambridge: Cambridge University Press.
- [3] Bybee, J. L. (1999). Usage-based phonology. In M. Darnell et al. (eds.), *Functionalism and Formalism in Linguistics, I: General Papers; II: Case Studies*. Amsterdam: Benjamins, pp. 211-242.
- [4] Bybee, J. L. (2000). The phonology of the lexicon: Evidence from lexical diffusion. In M. Barlow & S. Kemmer (Eds.), *Usage Based Models of Language* (pp. 65–86). Stanford: Center for the Study of Language and Information, pp. 65-86.
- [5] Bybee, J. L. (2001). *Phonology and language use*. Cambridge & New York: Cambridge University Press.
- [6] Donegan, P. J., & Stampe, D. (1979). The study of Natural Phonology. In Dan Dinnsen (Ed.), *Current approaches to phonological theory*. Bloomington: Indiana University Press, pp. 126-173.
- [7] Dziubalska-Kořaczyk, Katarzyna. 2007. Natural Phonology: Universal principles for the study of language (insiders meet outsiders). Paper presented at Special Session of 2007 ICPhS.
- [8] Jakobson, R. (1972). *Child language. Aphasia and phonological universals*. In *Janua linguarum; Series minor; nr. 72*; The Hague: Paris Mouton.
- [9] Lakoff, G. (1987). *Women, fire and dangerous things. What categories reveal about the mind*. Chicago: University of Chicago Press.
- [10] Langacker, R. W. (1987). *Foundations of Cognitive Grammar*. Stanford: Stanford University Press.
- [11] Langacker, R. W. (1995). Raising and transparency. *Language* 71(1):1-62.
- [12] Nathan, G. S. (1986). Phonemes as mental categories. In *Proceedings of the 12th Annual Meeting of the Berkeley Linguistics Society*. (Vol. 12, pp. 212–224).
- [13] Nathan, G. S. (1999). What functionalists can learn from formalists in phonology. In M. Darnell et al. (eds.), *Functionalism and Formalism in Linguistics, I: General Papers; II: Case Studies*. Amsterdam: Benjamins, (pp. 305–327).
- [14] Nathan, G. S. (2007). Is the Phoneme Usage-Based? – Some Issues. *International Journal of English Studies*, 6(2), 173–195. Murcia: Universidad de Murcia.
- [15] Pierrehumbert, J. (2002). Word specific phonetics. In C. Gussenhoven & N. Warner, *Laboratory Phonology 7*. The Hague: Mouton de Gruyter, pp. 101–139.
- [16] Rosch, E. (1975). Cognitive representations of semantic categories. *Journal of Experimental Psychology: General*, 104, 192–233.
- [17] Rosch, E. (1978). Principles of categorization. In Eleanor Rosch and B.B. Lloyd. (Ed.), *Cognition and Categorization*. Hillsdale, NJ: Lawrence Erlbaum, pp. 27–48.

- [18] Stampe, D. (1987). On phonological representation. In W. U. Dressler, H. C. Luschützky, O. Pfeiffer & J. R. Rennison (eds), *Phonologica 1984* (pp. 287–300). London: Cambridge University Press.
- [19] Steriade, D. (1996). Underspecification and markedness. In J. A. Goldsmith, *The handbook of phonological theory*. Cambridge, MA: Blackwell.
- [20] Treiman, R. (1993). *Beginning to spell: A study of first-grade children*. New York: Oxford University Press.