

---

## PHONEME COGNITION AND PERCEPTION

DON GEORGE\*

I used to think that in any language the phonemes were distinguished from each other on the basis of the physical characteristics of the sound spectrum and were identified by selected features which made each phoneme distinctive. Therefore, any deviation from the permitted allophonic variations in the language would shift the phoneme boundary and produce a significantly different utterance. Consequently, I devised numerous minimal-pair drills and loaded sentences in an attempt to condition the auditory and articulatory mechanisms of students to distinguish accurately those sound complexes to be cognized as the same from those to be cognized as different.

Meanwhile, I was a little bothered by the question of how we know when a person makes a phoneme substitution, or a non-permissible mutilation, which phoneme was intended. This ability, I argued, came as a result of phonetic training, and the argument became the rationale for courses in phonetics.

Some time ago in an advanced phonetics class, after I had been discussing some of the predictable phonetic problems of both native and non-native speakers and the importance of accurate auditory and articulatory drills to correct these "errors", an astute student asked, "Why is it that when a non-native speaker shifts phonemes or mutilates a phoneme even people with no phonetic training usually know what he is trying to say?" In other words, how does the phonetically untrained person cognize signals which are physically different as being the same, signals which are physically the same as being different, cognize missing elements in the signal and ignore excrescent elements?

After admitting the statistical necessity of controlling variables in experimental situations, and considering the various para-linguistic clues to meaning and the possible significance of the interaction of uncontrolled variables in normal communication, we attempted to arrive at an understandable answer to the question. The thesis which evolved may be stated somewhat as follows: Phoneme cognition is as much, if not more, dependent upon a complex of extra-phonetic clues as upon the physical characteristics of the acoustic signal. These clues may operate independently

---

\* University of Southern Mississippi.

or together, sequentially or simultaneously, and serve to validate normal cognition and to correct faulty acoustic signals.

If we take as a model for phoneme perception a series of monitors which regulate and classify the components of the total acoustic input, first censoring, or filtering out, any unwanted elements, and then, for the remainder, identifying the frequencies physically present, measuring the relative strength of the frequencies, and establishing the temporal relationships of the components, we have the basis for identifying the phonetic structure of an utterance whether it is cognized or not. This model, though, is inadequate to describe cognition. The precision with which the identification can be made is, of course, a matter of phonetic sophistication. Normal cognition—that is, speakers using the same code system—involves scanning the memory for lexicon, syntactical and morphological pattern, collocatability of items, situational probability, and adding to all these the evidence from other senses.

In spite of the large number of “emes” that have been coined since the term *phoneme* was introduced, I would like to add another to refer to the minimal element necessary to transmute a mutilated or missing acoustic signal into a cognizable unit. I would suggest the term *cogneme*.

As a simple illustration let me use a sentence in which at least three categories of cognitive correction are needed. Suppose one hears, “I cut a peak feesh.” One program consisting of the total lexical inventory scans the utterance and reports, “*Feesh* not in lexicon; [iy] = [i]. All lexical items now stand. But a second program of permitted collocations declares, “*Peak* not collocatable with *fish*.” Another scanning of the lexicon locates no probable compounds with *fish*, the stress pattern program indicates a qualifier, an inventory of possible qualifiers gives the minimal correction as, [p] = [b], [iy] = [i], and [k] = [g]. Evidence from other senses confirms the validity of *big fish*.

If a signal from the visual area reports a large, bloody knife, then *cut* is cognized as valid. If, on the other hand, a large, cleaned, wrapped fish is visually perceived the correction may be made, “Evidence for *cut* negative. Possession indicated. Change [k] to [g] and [ʌ] to [a].” Thus *cut* is cognized as *got*. But, if a freshly caught fish and fishing tackle is visually perceived, the report may come in as, “Evidence for *cut* negative. Possession negative. Change [ʌ] to [ɔ].” Thus, *cut* is cognized as *caught*.

Since this scanning is done almost instantaneously and performed sub-consciously, we are not aware of having done it, and our cognitive system hears the statement, “I caught a big fish.”

In addition to cognemes of lexicon, collocation, and perceived context already mentioned, there are those which deal in the same way with permitted versus not-permitted syntactical sequences and morphological structures of the type, “Me go now,” or “He give me it.” We cognize *me* as *I*, supply a missing auxiliary, shift a present to a past morpheme, and re-arrange the sequence of direct and indirect objects.

Though the perceived referent or the observed context may be all that is needed when the situation makes the referent obvious (as when a Bulgarian friend of mine said, “It’s nize here on your *teh-rahss*” when we were sitting on my terrace), when there is no visible referent or context, cognition will depend on a knowledge of possible versus not-possible referents. If more than one referent is possible, a tentative correction on the basis of probability will be made and a second cogneme will be elicited to confirm or negate the decision. If, for example, in a non-contextual situation you are asked, “You like docks?” either *dogs* or *ducks* may be considered equally probably, and the reply will likely be evasive until a subsequent cogneme establishes the referent more clearly.

Finally, such clues as vocal quality, facial expression, gesture, or some supra-segmental phoneme may serve as cognemes by revealing the intention of the speaker. Thus, when a Louisiana Cajun calls out, “Hey, you goddam dog?” even if we do not know he is searching for his lost hounds we know from one of these clues that he is not cursing us but is asking if we have his dogs.

In summary, though there is nothing profoundly new in these concepts which have been discussed by many writers, I have found that this model, a system of monitors processing the raw data of the physical stimulus, accepting, rejecting, substituting, adding, or taking from, according to the various levels of the total linguistic experience is a helpful device for explaining to students the complex way by which we cognize incomplete or inaccurate signals. This corrective ability permits a considerable degree of linguistic tolerance. Without it communication would be difficult, if not impossible, unless absolute phonemic conformity were achieved.