Proc. 5th int. Congr. phon. Sci., Münster 1964, pp. 231–234
(S. Karger, Basel/New York 1965).

Instituut voor Perceptie Onderzoek, Eindhoven

The Perception of Phonemes as a Function of Acoustic and Distributional Cues

By A. Cohen and A. F. V. van Katwijk, Eindhoven

Do phonemes have any kind of existence to the extent that they can be perceived? One may well argue that phonemes are the outcome of a linguistic operation on language material and may be described in terms of their distinctive character. It is by no means certain that this distinctive character has a perceptual correlate as such in the listener. On the contrary, there is substantial evidence that in speech perception a number of perceptual cues are operative that need not coincide with the distinctive features as postulated by *Jakobson* and *Halle*[1]. In fact, some phonemes can be recognized in isolation on the strength of inherent perceptual cues, such as colour and duration in the case of vowels.

That linguistic elements belonging to the same class of phenomena, in this case phonemes, should show a certain differentiation in the degree of autonomy need cause no surprise. On the morphemic level one generally distinguishes between free and bound forms. A similar observation may be made regarding the meaning of words. Some linguists hold that word meanings can be established only by studying the contexts in which these words are used. But if we ask: what is the meaning of the word "transubstantiation", do we necessarily force our informant to elicit a number of sentences in which the word may be used? On the other hand, there are words, e.g. functional words, which will put an informant in a quandary if he is not allowed to generate examples. We believe that much the same applies to the phonemes of a language and that they may be "known" by the language users either in isolation or in specific contexts.

To substantiate this belief one should try to create conditions in which language users are forced to make use of a type of knowledge that they are not explicitly aware of in normal speech situations.

Just as *Sapir*[2] was able to elicit from his analphabetic native informants the segmentation of whole utterances into single words, so all phoneticians work on the assumption that it is quite feasible to ask subjects, who may be phonetically naive, to respond to questions hinging on the perception of individual phonemes either in isolation or in context. We believe that this perception of individual phonemes is operative whenever a situation is created in which acoustic speech forms are presented divested of meaning. This is the case with PB lists, as used by communications engineers (who may be unaware of the linguistic implications of their work) in order to test the efficiency of a speech transmission channel. These involve e.g. CVC forms that contain segments, representing all the phonemes of a particular language in a way that the actual frequency of occurrence of separate phonemes is faithfully reflected in the total material. The whole burden of identification is therefore carried by the acoustic information contained in the nonsense forms.

So far these examples are largely concerned with the knowledge of phonemes by virtue of acoustic cues, but there are quite a number of data testifying to the existence of some kind of knowledge of the distributional rules of the occurrence of phonemes on the part of naive users of the language. Following *Shannon*[3] one may engage listeners in a guessing game by asking them to establish an utterance as a series of consecutive speech sounds while they are only told whenever they have made the right guess. It turns out that they show themselves well aware of the contextual constraints of their language by the decreasing number of guesses as more context becomes available. Evidence from a few more sources can be adduced to indicate knowledge in native speakers of word constituents devoid of meaning as such, as well as of the distributional rules prevailing among them: the acquisition of the mother tongue by children, where often words that are not acceptable have to be corrected mainly on an acoustic basis; a person solving a crossword puzzle or the occasional versifier on the look-out for rhyme words; evidence from diachronic linguistics and lapses in adult speech. Instances of sound shifts may be observed in everyday usage whenever native speakers are confronted with dialectal speech.

As for the lapses in speaking it may be observed that these can

be perceived either by the speaker, in which case he generally hastens to correct them, or by the listener acting as an observer on the look-out for cues about the general programming strategy applied in normal speech (see e.g. *E. H. Sturtevant*[4]). Such errors may lay themselves open to some kind of systematisation.

In short, the evidence of a language user's knowledge of linguistic elements, such as phonemes, is there for the asking. One may even go further and state that in some respects this knowledge obtrudes as is shown in the case of the obvious difficulties encountered by adult speakers when first confronted with a foreign language. Even though the individual phonemes are known, such as /s/ and /x/ in German, their unusual combination into a sequence such as in Dutch *Scheveningen* has traditionally been used as a highly efficient shibboleth to trap the unwary foreigner.

Now a rough classification can be made of the phonemes of a language as to their perceptibility on an acoustic or combined acoustic and distributional basis. Vowels stand out as basically identifiable on their own strength, with the exception of /ə/, as shown in experiments with synthetic speech. These experiments involved a firm knowledge on the side of the listeners of both spectral and temporal characteristics.

Consonants like /l, m, n, j, w/ which may occur in some cases only in conjunction with vowels are a lot harder to identify by themselves. It takes a tiny fraction of an adjoining vowel to bring about correct identification.

As for such consonants as plosives and fricatives which do not necessarily occur in languages of this type in direct contact with vowels, it seems doubtful if experiments limited to studying these consonants in this context can be considered conclusive.

In conclusion, phonemes may be seen on the one hand as the outcome of a purely linguistic analysis, on the other hand there seems to be ample reason for accepting them as entities, perceptually marked by a number of cues. The latter view implies that speech perception is hard to envisage as a unidimensional process in terms of distinctive features. These features are themselves an outcome of a phoneme analysis and have to be tested empirically before they can be accepted as perceptual cues.

References

1. *Jakobson, R.* and *Halle, M.:* Fundamentals of language; ch. I (Mouton, The Hague 1956).
2. *Sapir, E.:* Language; pp. 33–34 (Harcourt, Brace & Co., New York 1959).
3. *Shannon, C. E.:* Prediction and entropy of printed English. Bell Syst. Techn. J. *30:* 50–64 (1951).
4. *Sturtevant, E. H.:* An introduction to linguistic science; ch. IV, IX (Yale Un. Press, New Haven 1963).

Authors' address: A. Cohen, Technische Hogeschool, *Eindhoven* (The Netherlands).