Proc. 5th int. Congr. phon. Sci., Münster 1964, pp. 215–217
(S. Karger, Basel/New York 1965).

Predicting the Intelligibility of Words III

John W. Black, Columbus, Ohio

The degree and manner in which the intelligibility of a word relates to the intelligibility of the phonemes that comprise the word has been the topic of two investigations*. The first was of two parts. In each the written responses of 200 listeners were analyzed. One group had attempted to identify a set of approximately 600 words, about 1940; the other, a similar set, about 1950[1]. Both sets of words were heard through electronic communication systems by military personnel in the presence of high-level noise (108–110 db, resetting *C* on a General Radio sound level meter). In the earlier instance the listeners heard direct transmissions by the full complement of their 200 peers; in the later case, the listeners heard material that had been recorded by persons who were experienced in conducting researches in voice communication and were aware of the difficulties of communicating in noise. There was no overlap in speakers or listeners and minimal overlap in the words that were spoken.

A second study was conducted (1960) as an economical replication and refinement of either part of the first experiment and analysis[2]. Each of five workers in voice communication selected randomly twenty lists of twenty-five words from a 300-word master list. He recorded one list daily and then in the same laboratory period listened to the recordings made by his four colleagues. In this instance the listener, working from a voluntarily screened script of the speaker's words, exposed a written word subsequent to hearing it and wrote only his responses (receptions) that were in error. The economy in time is apparent. The refinement included the fact that all responses were made and, in turn, utilized. In the earlier instance there were many omissions among the responses; also single instances

* These studies and the present one were conducted at The Ohio State University, Columbus, Ohio, under a contract between the Office of Naval Research and The Ohio State University Research Foundation [Contract No. Nonr-495 (18) NR 145-993].

of particular responses – those made by one listener only. These were treated as bizarre and discarded.

The present experiment (1963) was in the main a replication of that portion of the foregoing one that fell subsequent to the recording stage. The same five listeners who heard the recordings of the earlier experiment listened a second time after a lapse of three years. In the present instance each listener could also hear the material that he had recorded earlier. This was not permitted in the original instance lest the effects of short-term memory contaminate the values of intelligibility. Additionally in the present study five new listeners provided a set of comparison values.

The rationale for the study lay particularly in reliability:

a) a comparison of relative intelligibility values of phonemes and words yielded by five listeners hearing the same verbal material a second time with three years intervening between the listening sessions, and

b) a comparison of the results yielded by a second group of listeners of the same limited size as the first.

Procedure

The recorded lists of one- and two-syllable words described above were heard by ten listeners at a O–db signal-to-noise ratio (white noise). Five of the listeners had recorded the words three years previously and each had listened to the recordings made by his colleagues; five were hearing them for the first time. The speakers had contributed phonetic transcriptions of their word lists at the time of recording. The listeners' phonetic transcriptions of the heard material were compared to these. This led to measures (a) of correct identifications of words (and concomitantly of the phonemes comprizing the words), (b) of correct identifications of phonemes in words that were misidentifications of a stimulus – called preservation-in-error (P–I–E) values, and (c) of the relative concurrence among substitutions of one phoneme for another.

The procedures for the former study utilized five methods of predicting the intelligibility of words from empirically derived values: 1. initial sound squared; 2. joint probability of the first two sounds, retaining the values peculiar to each position; 3. joint probability of the first two sounds, using the value for the initial position only; 4.–5. joint probability of the two and three most intelligible

sounds in the word. None of these appeared to hold an advantage over the recommendation of the earlier study, that the intelligibility of a word is related beyond chance to the joint intelligibility of the first two phonemes. Accordingly in the present study this method was employed.

Results

The three sets of data showed a significant correspondence among the relative intelligibility values of the various phonemes. The consonants varied in intelligibility from 60 to 90%; the vowels, 65 to 90%. The comparison group of listeners yielded the highest of the three sets of scores for each of one half of the consonants and for 10 of the 16 vowels and diphthongs.

On the basis of scores predicted from the joint intelligibility of the first two phonemes of a word, the predicted values would be the same for one- and two-syllable words. The obtained values of two-syllable words exceeded the values of one-syllable words. Thus a constant error is built into this method of predicting the intelligibility of words. The method accurately predicts the mean value of two-syllable words and over estimates the mean intelligibility value of one-syllable words.

Not only does the joint-intelligibility of the first two phonemes yield an acceptable mean value, it provides a distribution of scores that is comparable to the distribution of obtained scores. The standard deviations of the obtained scores in categories of at least 35 words ranged from 7.2 to 18.9; for the predicted values, 6.7 to 12.9.

Conclusions

The present data suggest that the relative intelligibility of a phoneme is a stable phenomenon; further that this characteristic can be demonstrated with few listeners, few speakers, and a limited number of words. Currently, this outcome is based on five speakers, five listeners, and 300 words, a combination that yielded 10,000 responses to simple words. Possibly more results are to be expected.

References

1. *Black, J. W.:* Predicting the intelligibility of words. Folia phoniat. *12:* 260–272 (1960).
2. *Black, J. W.:* Predicting the intelligibility of words II. Proc. Fourth Int. Congr. Phonetic Sciences; pp. 347–354 (Mouton, s'Gravenhage 1962).

Author's address: Prof. Dr. John W. Black, The Ohio State University, Department of Speech, 154 North Oval Drive, *Columbus 10*, Ohio (USA).