Published in final edited form as:
Lang Cogn Process. 1996 ; 11(6): 681–687.

Word Identification in Noise

David B. Pisoni
Indiana University, Bloomington, Indiana, USA

Abstract

Speech intelligibility has traditionally been measured by presenting words mixed in noise to listeners for identification at several different signal-to-noise ratios. The words are produced in isolation or in sentence contexts where the predictability of specific items can be varied. Psychometric functions are typically obtained relating signal-to-noise ratio to percent correct recognition. Error analyses are often carried out by examining response confusions to construct similarity spaces for words which reflect their perceptual organisation and acoustic–phonetic similarity. When using these techniques to measure speech discrimination or speech intelligibility in an open-set format, the recognition score obtained reflects the combined influence of both the sensory information encoded in the speech signal as well as the listener's decision process and response biases. Despite this limitation, the procedure has strong face validity as a measure of word recognition performance in normal-hearing listeners as well as other clinical populations which routinely use speech audiometry techniques to diagnose and assess both peripheral and central hearing impairments. All of the major findings and phenomena in the spoken word recognition literature can be demonstrated and explored with this experimental method. This technique continues to provide extremely valuable information about the organisation of words in the mental lexicon and how these sound patterns are accessed from acoustic-phonetic information in the speech signal.

Issues Addressed

1. Signal-to-noise (S/N) ratio needed for correct identification of words mixed in noise (Miller, Heise, & Lichten, 1951).
2. Underlying sensory and cognitive factors controlling intelligibility of words both in isolation and in various kinds of contexts (Miller et al., 1951; Kalikow, Stevens, & Elliott, 1977).

First Uses

Campbell (1910); Fletcher (1929); Egan (1948).

Description

Spoken words are mixed in noise at various S/N ratios and presented to listeners who are asked to recognise or identify the stimulus pattern as an English word (Egan, 1948; French & Steinberg, 1947). Sometimes nonsense syllables are used as well as pseudowords to dissociate early processes of speech perception, which are controlled primarily by the sensory information in the acoustic signal, from word recognition and lexical access, which also involve knowledge of the sound patterns and distinctive features of a particular

Requests for reprints should be addressed to David B. Pisoni, Speech Research Laboratory, Department of Psychology, Indiana University, Bloomington, IN 47405, USA. Pisoni@Indiana.edu.

language (Bagley, 1900; Cole & Rudnicky, 1983). The identification of nonsense syllables, words and pseudowords also requires that the listener have proper knowledge of the relevant phonological contrasts in a particular language.

Stimuli

Any linguistic stimulus of interest can be used, including syllables, words, pseudowords and sentences (Miller et al., 1951). In many studies, white noise is used to introduce degradation of the speech signal. However, other forms of stimulus degradation have been developed using envelope-shap ed noise, which provides a constant S/N ratio across consonant and vowels (Horii, House, & Hughes, 1971; O'Malley & Peterson, 1966).

Dependent Variables

1. Percent correct recognition (i.e. identification) using psychometric functions as a function of S/N ratio (Miller et al., 1951).
2. Confidence ratings (Pollack & Decker, 1958).
3. Response confusions in noise using error responses (Miller & Nicely, 1955; Wang & Bilger, 1973).
4. Scaling and construction of similarity spaces using MDS techniques (Shepard, 1972; Schiavetti, 1992; Triesman, 1978a).

Independent Variables

1. Signal-to-noise ratio.
2. White noise *vs* envelope-shaped noise.
3. Word frequency and familiarity.
4. Word length.
5. Lexical density (i.e. perceptual similarity).
6. Speaking rate.
7. Familiar *vs* unknown voices.
8. Sentence context.
9. Stimulus set size.
10. Auditory *vs* auditory + visual presentation.

Analysis Issues

Separation of sensory (acoustic-phonetic) properties from decision biases resulting from the use of top-down knowledge in perception.

Effects Found with Paradigm

1. Context

   Shown by: Miller et al. (1951); Kalikow et al. (1977); Huggins and Nickerson (1985); Miller (1962).

2. Word frequency and familiarity

Shown by: Owens (1961); Rosenzweig and Postman (1957); Savin (1963); Broadbent (1967).

3. Word length

   Shown by: Egan (1948).

4. Lexical neighbourhood

   Shown by: Triesman (1978b); Luce, Pisoni and Goldinger (1990).

5. Native *vs* non-native listeners

   Shown by: Lane (1963).

6. Normal *vs* hearing-impaired listeners

   Shown by: Hirsh et al. (1952); Penrod (1985).

7. Synthetic *vs* natural speech

   Shown by: Pisoni, Nusbaum, Luce and Slowiaczek (1985).

8. Differential masking of consonants and vowels

   Shown by: Licklider and Miller (1951); Hawkins and Stevens (1950); Miller (1947).

9. Voice familiarity effects

   Shown by: Peters (1955); Penrod (1979); Mullennix, Pisoni and Martin (1989).

10. Multi-modal audio-visual integration effects

    Shown by: Sumby and Pollack (1954).

11. Form-based priming effects

    Shown by: Slowiaczek, Nusbaum and Pisoni (1987).

Design Issues

1. Substantial set-size effects are found for digits, letters, nonsense syllables, words and sentences (Miller et al., 1951).
2. Substantial differences between open-set and closed-set response format demonstrates the important role of prior knowledge of response alternatives on spoken word recognition performance (Black, 1957; House, Williams, Hecker, & Kryter, 1965; Sumby, 1962).

Validity

There is strong “face validity” for this experimental paradigm (Hawley, 1977). Many, if not all, of the major phenomena in word recognition and spoken language processing can be demonstrated and studied experimentally using this method. The research literature on speech intelligibility is extensive, dating back well before the Second World War (Beranek, 1947; Black, 1946; Campbell, 1910; Egan, 1948; Fletcher, 1929; Licklider & Miller, 1951; Miller et al., 1951; Kalikow et al., 1977).

Advantages

1. Ease of use.

2. Permits the experimenter to control the amount of sensory information in the signal and degradation levels.
3. Permits the computation of psychometric functions for subjects' identification data as a function of the S/N ratio.
4. Provides a way to examine the underlying perceptual and psychological processes used in recognising words from degraded or partial information.
5. Permits the use of a number of different dependent variables which provide converging evidence on processes of word recognition and lexical access.

Potential Artifacts

The only potential artifacts are that the observed recognition score (percent correct recognition) reflects the combined use of both sensory information in the signal—that is, bottom-up acoustic-phonetic processing—as well as top-down lexical processing, reflecting the contribution of knowledge of the listener based on the sound patterns in his or her language that are stored in long-term lexical memory. White noise also masks consonants more than vowels. And there are differences in intelligibility among different talkers and among different words.

Problems

There are several problems with this technique. The data on masking of the acoustic–phonetic properties of speech (Miller & Nicely, 1955) are still largely empirical and there is no current model that combines knowledge of acoustics–phonetics with masking theory to accurately predict speech recognition performance at this early level of perceptual analysis. Because of the open-set nature of the word recognition task, it has been extremely difficult to separate the contribution of sensory information in the signal from the listener's prior knowledge and response biases that arise in the decision process.

Uses with Other Populations

In addition to its traditional use in the assessment of telephone and communication equipment with normal-hearing listeners, which has a very long history, speech intelligibility techiques have also become routine in the speech clinic to assess and diagnose a wide range of hearing and speech perception disorders in clinical populations. Known in this context as “clinical speech audiometry” or just “speech discrimination” tests, the same stimulus materials and methods have been used to measure word recognition performance in both the quiet and in noise (see Hirsh et al., 1952; Hudgins, Hawkins, Karlin, & Stevens, 1947; Davis & Silverman, 1947). As commonly used, the term “speech intelligibility” refers to the reproduction of speech by a transmission system, whereas the term “speech discrimination” is used more routinely in audiology for the clinical assessment of a human listener's ability to perceive and understand speech (Penrod, 1985; Owens & Schubert, 1968; Schubert & Owens, 1971).

Other Comments

Speech intelligibility tests have also been used to study multi-modal integration of auditory and visual information (Sumby & Pollack, 1954) and to place the problem of speech perception within the larger context of event perception and recent developments in ecological psychology (Fowler, 1986; Gaver, 1993).

Acknowledgments

Preparation of this paper was supported by NIH Research Grant DC-00111 to Indiana University in Bloomington.

References

Bagley WC. The apperception of the spoken sentence: A study in the psychology of language. American Journal of Psychology. 1900; 12:80–130.

Beranek LL. The design of speech communication systems. IRE Proceedings. 1947; 35:880–890.

Black JW. Studies in speech intelligibility: A program of war-time research. Speech Monographs. 1946; 2:1–68.

Black JW. Multiple choice intelligibility tests. Journal of Speech and Hearing Disorders. 1957; 22:213–235. [PubMed: 13429664]

Broadbent DE. Word-frequency effect and response bias. Psychological Review. 1967; 74:1–15. [PubMed: 5341440]

Campbell GA. Telephonic intelligibility. Philosophical Magazine. 1910; 19:152–159.

Cole RA, Rudnicky AI. What's new in speech perception? The research and ideas of William Chandler Bagley. Psychological Review. 1983; 90:94–101. [PubMed: 6342015]

Davis, H.; Silverman, SR. Hearing and deafness. Holt, Rinehart and Winston; New York: 1947.

Egan JP. Articulation testing methods. Laryngoscope. 1948; 58:955–991. [PubMed: 18887435]

Fletcher, H. Speech and hearing. Van Nostrand; New York: 1929.

Fowler CA. An event approach to the study of speech perception from a direct-realist perspective. Journal of Phonetics. 1986; 14:3–28.

French NR, Steinberg JC. Factors governing the intelligibility of speech sounds. Journal of the Acoustical Society of America. 1947; 19:90–119.

Gaver WW. What in the world do we hear? An ecological approach to auditory event perception. Ecological Psychology. 1993; 5:1–29.

Hawkins JE Jr, Stevens SS. The masking of pure tones and of speech by white noise. Journal of the Acoustical Society of America. 1950; 22:6–13.

Hawley, M., editor. Benchmark papers in acoustics, Vol. 11: Speech intelligibility and speaker recognition. Dowden, Hutchinson and Ross; Stroudsburg, PA: 1977.

Hirsh IJ, Davis H, Silverman SR, Reynolds EG, Eldert E, Benson RW. Development of materials for speech audiometry. Journal of Speech and Hearing Disorders. 1952; 17:321–337. [PubMed: 13053556]

Horii Y, House AS, Hughes GW. A masking noise with speech envelope characteristics for studying intelligibility. Journal of the Acoustical Society of America. 1971; 49:1849–1856. [PubMed: 5125732]

House AS, Williams CE, Hecker MH, Kryter KD. Articulation testing methods: Consonantal differentiation with a closed response set. Journal of the Acoustical Society of America. 1965; 37:158–166. [PubMed: 14265103]

Hudgins CV, Hawkins JE, Karlin JE, Stevens SS. The development of recorded auditory tests for measuring hearing loss for speech. Laryngoscope. 1947; 57:57–89. [PubMed: 20287775]

Huggins AW, Nickerson RS. Speech quality evaluation using phoneme-specific sentences. Journal of the Acoustical Society of America. 1985; 77:1896–1906. [PubMed: 3998299]

Kalikow DN, Stevens KN, Elliott LL. Development of a test of speech intelligibility in noise using sentence materials with controlled word predictability. Journal of the Acoustical Society of America. 1977; 61:1337–1351. [PubMed: 881487]

Lane H. Foreign accent and speech distortion. Journal of the Acoustical Society of America. 1963; 35:451–453.

Licklider, JCR.; Miller, GA. The perception of speech. In: Stevens, SS., editor. Handbook of experimental psychology. John Wiley; New York: 1951. p. 1040-1074.

Luce, PA.; Pisoni, DB.; Goldinger, SD. Similarity neighbourhoods of spoken words. In: Altmann, G., editor. Cognitive models of speech processing: Psycholinguistic and computation perspectives. MIT Press; Cambridge, MA: 1990. p. 122-147.

Miller GA. The masking of speech. Psychological Bulletin. 1947; 44:105–129. [PubMed: 20288932]

Miller GA. Decision units in the perception of speech. IRE Transactions on Information Theory. 1962; IT-8:81–83.

Miller GA, Nicely PE. An analysis of perceptual confusions among some English consonants. Journal of the Acoustical Society of America. 1955; 27:338–352.

Miller GA, Heise GA, Lichten W. The intelligibility of speech as a function of the context of the test materials. Journal of Experimental Psychology. 1951; 41:329–335. [PubMed: 14861384]

Mullennix JW, Pisoni DB, Martin CS. Some effects of talker variability on spoken word recognition. Journal of the Acoustical Society of America. 1989; 85:365–378. [PubMed: 2921419]

O'Malley MH, Peterson GE. An experimental method for prosodic analysis. Phonetica. 1966; 15:1–13.

Owens E. Intelligibility of words varying in familiarity. Journal of Speech and Hearing Research. 1961; 4:113–129. [PubMed: 13731816]

Owens E, Schubert ED. The development of consonant items for speech discrimination testing. Journal of Speech and Hearing Research. 1968; 11:656–667. [PubMed: 5722493]

Penrod JP. Talker effects on word-discrimination scores of adults with sensorineural hearing impairment. Journal of Speech and Hearing Disorders. 1979; 44:340–349. [PubMed: 480938]

Penrod, JP. Speech discrimination testing. In: Katz, J., editor. Handbook of clinical audiology. 3rd edn. Williams and Wilkins; Baltimore, MD: 1985. p. 235-255.

Peters, RW. Joint Report No. 56, U.S. Naval School of Aviation Medicine. Pensacola, FL: 1955. The relative intelligibility of single-voice and multiple-voice messages under various condition of noise; p. 1-9.

Pisoni DB, Nusbaum HC, Luce PA, Slowiaczek LM. Speech perception, word recognition, and the structure of the lexicon. Speech Communication. 1985; 4:75–95.

Pollack I, Decker LR. Confidence ratings, message reception, and the receiver operating characteristic. Journal of the Acoustical Society of America. 1958; 30:286–292.

Rosenzweig MR, Postman L. Intelligibility as a function of frequency of usage. Journal of Experimental Psychology. 1957; 54:412–422. [PubMed: 13491767]

Savin HB. Word-frequency effect and errors in the perception of speech. Journal of the Acoustical Society of America. 1963; 35:200–206.

Schiavetti, N. Scaling procedures for the measurement of speech intelligibility. In: Kent, RD., editor. Intelligibility in speech disorders. John Benjamins; Amsterdam: 1992. p. 11-34.

Schubert ED, Owens E. CVC words as test items. Journal of Auditory Research. 1971; 11:88–100.

Shepard, RN. Psychological representation of speech sounds. In: David, EE.; Denes, PB., editors. Human communication: A unified view. McGraw-Hill; New York: 1972. p. 67-113.

Slowiaczek LM, Nusbaum HC, Pisoni DB. Phonological priming in auditory word recognition. Journal of Experimental Psychology: Learning, Memory and Cognition. 1987; 13:64–75.

Sumby WH. On the choice of strategies in the identification of spoken words mixed with noise. Language and Speech. 1962; 5:119–124.

Sumby WH, Pollack I. Visual contribution to speech intelligibility in noise. Journal of the Acoustical Society of America. 1954; 26:212–215.

Triesman M. A theory of the identification of complex stimuli with an application to word recognition. Psychological Review. 1978a; 85:525–570.

Triesman M. Space or lexicon? The word frequency effect and the error response frequency effect. Journal of Verbal Learning and Verbal Behavior. 1978b; 17:37–59.

Wang MD, Bilger RC. Consonant confusions in noise: A study of perceptual features. Journal of the Acoustical Society of America. 1973; 54:1248–1266. [PubMed: 4765809]