PHONETICALLY BASED NEW METHOD FOR AUDIOMETRY: THE G-O-H MEASURING SYSTEM USING SYNTHETIC SPEECH

MÁRIA GÓSY GÁBOR OLASZY JENŐ HIRSCHBERG ZSOLT FARKAS

Dept. of Phonetics
Linguistics Institute
Budapest 1250 Pf. 19. Hungary

Heim Pál Children's
Hospital
Budapest 1089 Hungary

ABSTRACT

A special phonetically based method in generating artificial monosyllabic words was developed for audiometric measurements. Using this synthesized material, the hearing capacity and the speech perception and understanding level of children can be judged easily. It is very important to examine regularly the hearing capacity of children and the evaluation of speech understanding level from the point of view of learning writing and reading in school. The speciality of our artificial words is in their low-redundancy acoustic structure and in the special frequency structure.

INTRODUCTION

There is a close connection between the articulation and perception bases of the process of speech acquisition. The initial development of perception and understanding abilities precedes that of speech production, but this difference between them subsequently decreases: their further development is assumed to take place in a permanent interaction. The bases of speech perception and understanding is hearing; this does not mean, however, that good hearing automatically ensures the normal processing of speech perception/understanding. That is why regular examination of hearing and understanding is very important, particularly in the early years when acquisition of the mother tongue is in progress. The identification of speech production problems is easier than that of speech perception/understanding ones. The normal communication situations provide a better opportunity for adults to detect the speech errors of children, revealing articulatory or grammatical problems. However, perception and/or understanding/comprehension difficulties can remain hidden because of various supplementary and compensatory strategies of children. This fact leads to delayed diagnosis and to difficulties in carrying out the appropriate corrective procedures. There are a lot of well-known problems related to the hearing measurements and mass screening of children between the ages of 3 and 7. What are the criteria that a suitable method for auditory screening of these small children has to meet? First: the signal that is given to the child's ear should be natural and familiar for him. Second: the measuring task should be easy to understand, that is, we should make it easy for the child to understand what he has to do during examination. Third: the measuring method should yield the highest possible amount of information about the hearing mechanism operative between 200-8000 Hz. These expectations are all met by our new screening procedure: the G-O-H system.

METHOD AND MATERIAL

On the basis of the results of a perceptual examination of Hungarian speech sounds whereby the values of their invariant cues are determined [1], the process of speech understanding can be further studied: the hearing mechanism and the level of recognition of words can be examined. The examination of the two processes can be combined if we produce speech material which only involves acoustic values corresponding to invariant features (or hardly more than that). This condition is satisfied by computer-generated, artificial speech based on perceptual data.
Speech as an auditory stimulus is familiar for children, and repeating sound-sequences is a natural task when the child acquires his first language, and repeats the words of his mother. Human speech is suitable for judgement of understanding level, but the speech-audiometric results cannot give exact data about the hearing capacity or about the extent and type of impairement, because natural speech is very redundant as to its frequency structure. The redundancy of speech means that speech sounds contain far more building elements than would be necessary for understanding. That is why natural speech can be understood in the case of certain hearing impairements: the redundant elements give an opportunity to guess the meaning. Our specially synthesized words contain only the necessary frequency components of each sound. The difference between the natural and synthesized words is only in the redundancy of the frequency structure (cf. Fig. 1). In spite of this difference they sound very similar. The speech perception threshold proved to be the same as the normal one obtained by using natural words.
How does the screening function of the artificially produced words work? Let us suppose that the system of speech understanding has to analyse data of quantity x to understand the word bus. But the acoustic

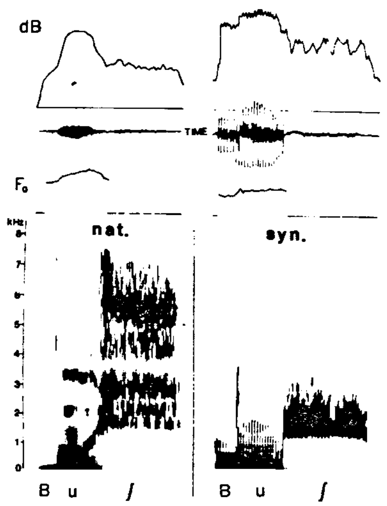

Fig. 1. Acoustic structures of natural and synthesized German word Busch

structure of natural speech is highly redundant, i.e. it contains significantly more information (data) in the speech than is necessary for its safe recognition. But the word bus contains data of quantity $x+y$. The data surplus (y) becomes stored and can be immediately called out in case of any kind of 'disorder' (e.g. noise), to provide supplementary information. On the other hand the word bus we synthesize hardly contains more information than the necessary quantity x. Therefore, in case there is some 'disorder' at any point in the recognition process (hearing loss, central problem), $x+y$ would make identification possible, but x itself does not, where comprehension will be mistaken (to some extent). The comprehension of a signal sequence containing information x requires the processing of all information in a perfectly sound fashion, e.g. by the help of normal hearing.
To provide a basis for the G-O-H method, a special test material was constructed which involved 44 meaningful monosyllabic Hungarian words. The criteria for choosing the words were as follows: (i) the monosyllabic words should have two or three speech sounds without consonant clusters, (ii) the words should contain speech sounds where the frequency parameters seem to serve as an acoustic cue for their identification, (iii) the test material includes three types of words containing only high-frequency sounds (like [sy:z]; words containing only low-frequency sounds (like [bu:]); words containing both high and low frequency sounds (like [bus]); (iv) an effort was made to collect a material exhibiting almost all Hungarian speech sounds in different sound-contexts and phonetic positions, (v) most of the words should be familiar for children of ages 3-7, however, the sample should also include a few items that are meaningless sound-sequences for the children. Attention was also payed to the order of the words: low-frequency and high-frequency words alternate with one another. So all children have an experience of success, because they can understand and repeat correctly at least every second word. This is important for good co-operation between the child and the examiner.
The three frequencies (500, 1000, 4000 Hz) used in pure tone audiometry seem to be very insufficient for the evaluation of speech understanding level. In normal hearing the acoustic information received at these frequencies accounts for some 60% of understanding. This means that, in cases of hearing losses at other frequencies, the child - screened and judged to have normal hearing - cannot understand speech correctly. Our G-O-H method solves this problem as well. Experiments were carried out with the G-O-H test material in clinic and day-care-centers with the participation of 400 normal-hearing and 150 hearing impaired children. People with normal hearing understand both human speech and the special synthesized artificial words equally well. But the hearing impaired patients cannot correctly understand the synthesized words because of the lack of redundant building elements.

RESULTS

Speech synthesis gives us an opportunity to define the desired frequency bands in speech sounds. These facts lead to the perceptual/understanding differences between the normal hearing and impaired hearing listeners. For example, a high frequency hearing impaired child with hearing loss above 5000 Hz cannot understand the word [se:l] 'wind' if the noise component of the initial consonant of the word is above 5000 Hz. In this case, the child receives acoustic information that he identifies as consonant like [f,h] or [t] depending on the extent of the hearing loss of the child. In the case of a slight hearing loss above 5000 Hz, the child will identify the sound-sequence as [fe:l] 'he is afraid of sg' which is an existing Hungarian word. In the case of somewhat more severe loss of hearing above 5000 Hz, the child will identify the sound-sequence as [he:l] which has no meaning in Hungarian, and with more severe loss he will identify, instead of the spirant [s], a stop consonant like [t,p] or [k]. In the case mentioned, the child identifies the word as [te:l] 'winter', because it is a frequent item in children's vocabulary (Figure 2 shows a German example). From the answers of the listeners' judgements can be made about the extent of their hearing losses. The answers are regular consequences of hearing losses, they are not results of imagination. (Experiments with filtered words confirmed us about these regular changes in perception.) Mass-measurements were carried out together with a control pure-tone audiometry corresponded to the results gained by the G-O-H method. This method, however, gave information about the understanding level as well. (In some cases the child did not co-operate in pure-tone audiometric examination, but he repeated the artificial words of the G-O-H device, so his hearing could be measured.)

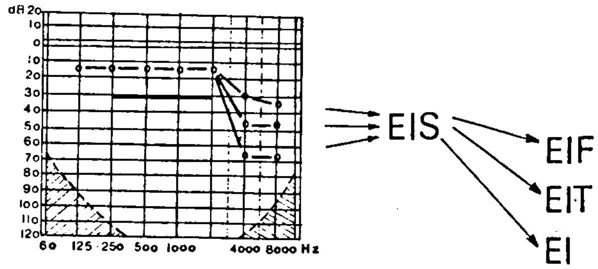

Fig. 2. Changes of possible responses in the case of different hearing losses

The possible answers of normal-hearing and impaired listeners were predicted theoretically on the basis of perceptual investigations concerning the acoustic cues of Hungarian speech sounds. Then, laboratory and clinical experiments were carried out, and the theoretically established types were slightly revised on the basis of the data obtained. Finally, the possible answers were arranged on an answer sheet according to the degree of hearing losses (Table 1).

Twenty monosyllabic German words have been developed so far in our laboratory. These synthesized words are suitable for application the G-O-H method in German too. Experiments were carried out with German-speaking children in kindergartens in Austria.

For the everyday use of this procedure, a measuring set has been developed. This portable case contains a playing system, a tape with the artificial words, an amplifier, a headphone and answer sheets. More than 150 of these sets with G-O-H system are being used in Hungary. To carry out the examinations there is no need for any expert, physician, phonetician or audiologist; it can be used by nurses, kindergarten teachers, speech therapists and so on.

The G-O-H method is a good tool for (i) finding out whether the child is mature enough to acquire writing and reading, (ii) detecting hearing problems, (iii) learning if the child with speech errors has perceptual problems too or not, (iv) detecting dyslexia, i.e. the disturbance in writing and reading.

Table 1. The Hungarian answer sheet (part) of the G-O-H system for measuring the hearing capacity from 200 Hz to 8000 Hz

Sor-szám	I. Jó hallás	II. Hallás- vagy beszédmegértési zavar állhat fenn, esetleg figyelmetlenség. A vizsgálat megismétlendő.	III. Halláoccupied zavar valószínű, orvosi vizsgálat javasolt.	IV. Hallászavar biztosra vehető, mielőbbi klinikai vizsgálat szükséges.
1.	meggy (megy)	begy legy vegy negy	egy	bó od e ó –
2.	síp (sík)	sít sűt sűp szíp szép	zúg suk sut su só víd fut hó	kút út tú ú –
3.	bú	dú bók but bot bó pók pú púk dú gú	tú tó pú pó út	ó ú –
4.	ász	ház pász áz	ás ágy áll áj	áf áh át áp á a ó ú –

The possible answer categories are the following: I. good hearing; II. slight hearing problem; III. hearing problem is probable; IV. hearing problem is certain, urgent clinical examination is required.

REFERENCE

[1] Gósy, Mária: Magyar beszédhangok felismerése, a kísérleti eredmények gyakorlati alkalmazása. Magyar Fonetikai Füzetek 15. 1986, 3-100.

Se 66.4.4