Lund University, Dept. of Linguistics
Working Papers 36 (1990), 99-108

On the Perception of Mood in Speech: Implications for the Hearing Impaired.

David House

Abstract

This paper presents an outline of a newly initiated project on the perception and production of mood by the hearing impaired. A preliminary perception experiment is reported where various manipulations of a semantically neutral utterance recorded in three different moods are tested. Results showed that hearing-impaired listeners have considerable difficulty in identifying speaker mood when compared to listeners with normal hearing. It is proposed that an interplay of several different acoustic cues is necessary for an unambiguous perception of mood and that this interplay is often not accessible to the hearing-impaired listener.

INTRODUCTION

As part of her extensive work on intonation, Eva Gårding has investigated the relationship between intonation and the production and perception of mood in American English (Gårding & Abramson 1965) and in a cross-linguistic study of intonation (Gårding 1986). In these studies she demonstrates the importance and universality of different contours as major cues used to signal different speaker moods and attitudes. For the hearing-impaired listener, however, these cues may not always be accessible or sufficient to unambiguously perceive the mood and attitude of the speaker.

Öster & Risberg 1986 showed that hearing-impaired children and adults have difficulty in identifying the mood of speakers in test sentences, particularly confusing happy with angry and sad. Similar confusions were obtained from normal listeners when presented with low-pass filtered test sentences. Hearing-impaired speakers may also experience a similar insecurity when expressing their own moods and attitudes in speech.

This paper represents a report from a newly initiated project concerning speech communication and the hearing impaired, specifically the perception and production of mood. The project will be carried out by Gösta Bruce, David House and Ursula Willstedt and is sponsored by the Bank of Sweden Tercentenary Foundation.

The aim of the project is to investigate the acoustic cues which can be used to signal speaker mood and the influence of hearing impairments on

the perception of these cues. By analyzing different moods produced by hearing-impaired speakers, we intend to relate the perception of the major acoustic cues by the hearing-impaired listeners to their own production of these cues. This analysis will form the basis for the development and testing of a speech therapy program designed to increase the awareness of how different moods are signaled in speech. The final goal of both the therapy program and the project is to improve the perception and production of mood in speech by the hearing impaired including cochlear implant patients. The project can be summarized by the following five steps: Perception-Production-Therapy-Production-Perception. See House 1990 for a preliminary project report.

Apart from the previously mentioned study by Öster & Risberg 1986 few investigations have been made concerning the perception of mood by hearing-impaired listeners. Earlier studies have mainly dealt with the perception of fundamental frequency cues and intonation. Risberg & Agelfors 1978 found that hearing-impaired subjects had difficulties in using fundamental frequency information to identify the emphasized word in a sentence. Fourcin 1980 and Engen et al. 1983 found similar difficulties in the perception of intonation by hearing-impaired children.

Previous studies of the production and perception of mood and attitude (e.g. Hadding-Kock 1961, Gårding & Abramson 1965, Williams & Stevens 1972, Fónagy 1981, Bruce 1982) are also in general agreement that fundamental frequency information (range and mean level) is the most important cue to speaker mood. It is clear, however, that other features such as intensity, tempo, voice quality and spectral configuration can play an important role in signaling speaker mood.

PRELIMINARY PERCEPTION TEST

Stimulus design

In an attempt to investigate the relative saliency of different acoustic cues in the perception of mood for both normal-hearing and hearing-impaired listeners, a recording of the semantically neutral number 2510 (in Swedish) (see Gårding & Abramson 1965) in three different moods (angry, happy and neutral) was used as a basis for a preliminary perception test. The recording was originally made by Gårding 1986 and used in an investigation of intonation and mood in several languages. From the original utterances, three additional versions were made: a low-pass filtered version (300 Hz, -48dB/oct.), a version in which the fundamental frequency

contour was stylized using a linear interpolation method (House et al. 1988), and a manipulated version in which the stylized Fo contours for angry and neutral were interchanged by means of LPC synthesis. The original utterances plus the additional three versions comprised a total of 11 stimuli.

A test tape was made in which each stimlus appeared five times. The test was divided into four parts. The first part comprised the filtered version only. The second and fourth parts comprised both original and stylized versions while the third part comprised the interchanged versions (interchanged Fo contours for angry and neutral). Since there were only two interchanged stimuli, five tokens of the stylized happy version were also included in part three as distractors. The stimuli were randomized within each part and presented with a five-second pause between stimuli. Four practice stimuli (filtered version) were presented before part one.

Subjects

Eleven subjects (16 to 75 years of age) with varying degrees of sensorineural hearing losses participated in the experiment. Eleven normal-hearing subjects with a similar age distribution comprised the control group.

Test procedure

The test was administered separately to each subject via a Revox A77 tape recorder and a high quality loudspeaker. Each subject was given a printed instruction sheet and an answer sheet with four mood options: angry, happy, sad and neutral. Sad was included as preliminary testing indicated that the neutral version could also be interpreted as sad as could one of the interchanged versions. The subjects were, however, permitted to freely write down any other mood they thought applied. The test took about 15 minutes including instructions and a pause between each of the four parts.

Results

The results of the original and filtered versions replicated the Öster & Risberg results (see figure 1). The hearing-impaired group had special difficulty perceiving the happy mood in the original version while the normal-hearing listeners had the same difficulty with the filtered version.

A=Angry, H=Happy, S=Sad, N=Neutral

NORMAL-HEARING LISTENERS n=11

Response in %

ORIGINAL STIMULI

Intended category	A	H	S	N
A	94	2	2	2
H	-	87	-	13
N	-	-	49	51

LP-FILTERED STIMULI, 300 Hz, -48dB/oct.

Intended category	A	H	S	N
A	76	13	5	5
H	9	29	9	53
N	-	-	42	56

STYLIZED STIMULI (LINEAR PITCH MOVEMENT)

Intended category	A	H	S	N
A	91	9	-	-
H	4	82	5	9
N	-	2	24	74

INTERCHANGED PITCH STIMULI (stylized)

Intended category	A	H	S	N
N-A	2	20	40	38
A-N	65	15	11	9

HEARING-IMPAIRED LISTENERS n=11

Response in %

ORIGINAL STIMULI

Intended category	A	H	S	N
A	85	11	2	-
H	7	56	11	24
N	-	7	25	67

LP-FILTERED STIMULI, 300 Hz, -48dB/oct.

Intended category	A	H	S	N
A	55	20	15	7
H	11	13	27	49
N	7	4	44	42

STYLIZED STIMULI (LINEAR PITCH MOVEMENT)

Intended category	A	H	S	N
A	87	4	9	-
H	5	38	29	24
N	2	5	33	60

INTERCHANGED PITCH STIMULI (stylized)

Intended category	A	H	S	N
N-A	-	15	42	42
A-N	53	22	7	15

Figure 1. Confusion matrices for four versions of the preliminary perception test, two listener groups. N-A = neutral stimulus with the 'angry' Fo contour, A-N = angry stimulus with the 'neutral' contour.

As expected, the neutral version was judged as ambiguous between sad and neutral by both groups in both versions.

Results for the stylized versions of the stimuli for the normal-hearing group were very similar to those of the original stimuli (figure 1). For the hearing-impaired group, however, identification of the happy mood deteriorated.

The interchanged versions showed considerable variation for both groups with a tendency for an 'angry' Fo contour on 'neutral' segmentals to be perceived as sad while a 'neutral' Fo on 'angry' segmentals was perceived as angry. In no part of the test did any subject choose a mood category other than the four suggested.

Discussion

These preliminary results reconfirm the results of the Öster & Risberg study showing that the perception of mood in speech can be a problem for the hearing impaired. The results further indicate that while fundamental frequency information is an important cue to the perception of mood, the perception of this cue alone is not sufficient to unambiguously perceive mood.

The elimination of many high-frequency spectral cues by low-pass filtering had a particularly detrimental effect on the identification of the happy utterance by both listener groups. Fundamental frequency stylization of the happy utterance, however, mainly affected the hearing-impaired group. This seems to indicate that the hearing-impaired group made use of fundamental frequency information as a cue to the happy mood, information which was partially lost during stylization while the normal-hearing group was able to use other (e.g. spectral) information which was retained even after stylization.

The confusions resulting from the interchanged stimuli demonstrate the importance of the interplay between different acoustic cues for both groups. We can propose that the salient cues to the perception of mood probably lie in the interplay between fundamental frequency, intensity dynamics, spectral characteristics, voice quality and durational aspects.

BASIC MOOD TEST

To further investigate the interplay of acoustic cues, we are designing a basic mood test using four intended moods: angry, happy, sad and neutral and two different semantically neutral sentences with situations described

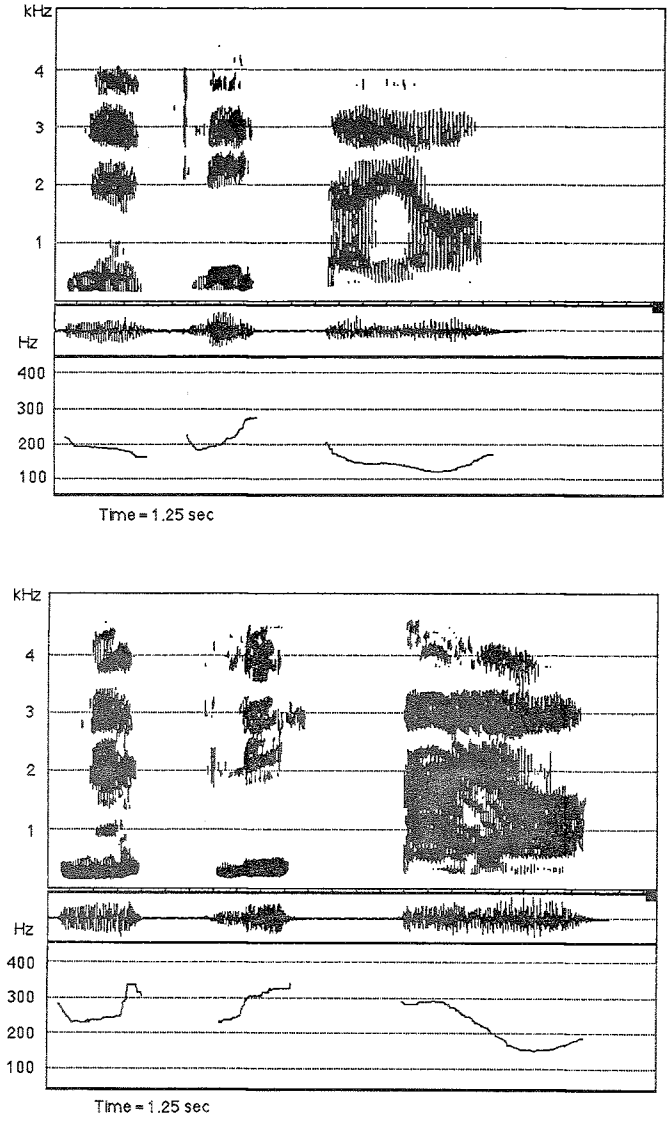

Figure 2. Spectrogram, waveform and fundamental frequency contour for the utterance, 'Nu flyttar jag' (Now I'm going to move): neutral version (top panel) and happy version (bottom panel).

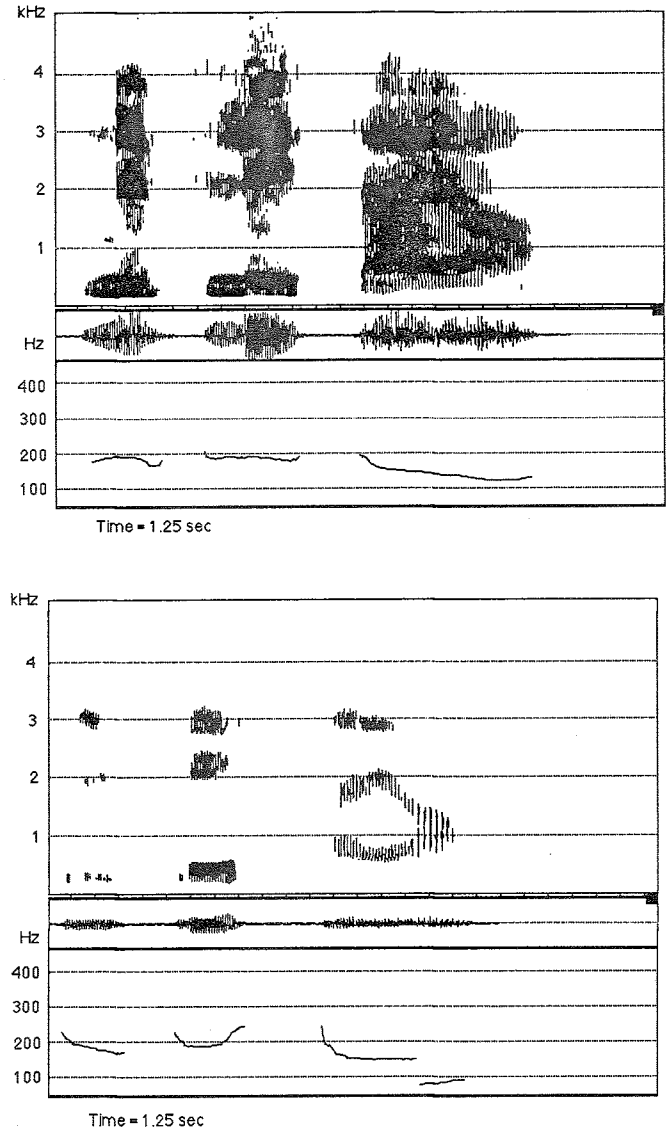

Figure 3. Spectrogram, waveform and fundamental frequency contour for the utterance, 'Nu flyttar jag' (Now I'm going to move): angry version (top panel) and sad version (bottom panel).

which correspond to the four moods. This test will be used both to elicit production of moods and to test perception of moods.

The two sentences are the number 2510 as in the preliminary test and the sentence 'Nu flyttar jag' (Now I'm going to move). For the number 2510 the situations are as follows: NEUTRAL – the number of books in the library, HAPPY – the sum of an unexpected tax rebate, ANGRY – the sum of an unexpected tax debt, SAD – the loss of a purse containing the unexpected tax rebate. For the sentence 'Nu flyttar jag' (Now I'm going to move) the following situations are described: NEUTRAL – reporting a change of address to an insurance company, HAPPY – telling friends about the acquisition of a great new apartment, ANGRY – telling friends about the landlord who has drastically raised the rent for an apartment in disrepair, and SAD – telling friends the news of a separation and its consequences.

The test has been used to elicit productions of moods by four cochlear implant patients (Nucleus multichannel device, see e.g. Waltzman & Hochberg 1990), both before and after the implant operation. Additional recordings are planned at six-month intervals. At least four additional implant patients are expected to participate in the production study. Although the recordings have not yet been studied in detail a tendency toward increased mood separation can be observed in the post-operative recordings.

For the perception study, the two utterances were recorded in the four different moods by a female speaker of southern Swedish. The speaker is a trained speech pathologist. Spectrograms, waveforms and fundamental frequency contours of the sentence 'Nu flyttar jag' in the four moods are presented in figures 2 - 3.

The neutral version (figure 2, top) is useful as a reference for comparing the other versions and can also be used as a point of departure for parameter manipulations. Compared to the neutral version, the happy version (figure 2, bottom) shows a greater Fo range with higher maximum values and a greater overall intensity as can be seen in the waveform and the spectrogram. There is also evidence of a raising of F4 compared to all other moods. This could be a result of a shortened oral tract due to lip spreading (smiling).

The angry version (figure 3, top) shows a compressed Fo range with lower maximum values than both the neutral and happy versions. This compressed Fo range coupled with the even greater overall intensity than in

the happy version may be an important cue to the angry mood, although anger can also be signaled by a large Fo range more similar to that of the happy mood.

Finally, the sad version (figure 3, bottom) demonstrates a particularly low overall intensity compared to all the other versions. The Fo contour is similar to that of the neutral version although the maximum value is lower. Also apparent is final creaky voice visible in the spectrogram and in the Fo contour. Creaky voice may also be an important signal of sadness.

A perception test was constructed using the eight utterances. The test has been tried out on two cochlear-implant patients resulting in identification confusions similar to those obtained on our preliminary test.

Our intention is to use the basic mood test as a point of departure for the manipulation of acoustic parameters to see which cue or combination of cues is most useful for and accessible to hearing-impaired listeners in facilitating the task of perceiving mood in speech.

CONCLUSIONS

Results from the preliminary perception test indicate that hearing-impaired listeners can have considerable difficulty in perceiving the mood of the speaker. Results from the filtered, stylized and manipulated version of the test also indicate that the cues to the perception of mood in speech probably lie in the interplay between several acoustic cues such as fundamental frequency, intensity dynamics, spectral chatracteristics, voice quality and durational aspects. Furthermore, the use of these cues by the hearing impaired may differ considerably from that of normal-hearing listeners. By developing a basic mood test we hope to create a useful tool for the analysis of both production and perception of mood by the hearing impaired including cochlear implant patients.

ACKNOWLEDGEMENTS

We are grateful to Eva Gårding whose ideas helped to inspire this project and to the Department of Otorhinolaryngology, University Hospital, Lund, for making available both patients and equipment. We would also like to thank Karin Andersson for letting us record her 'moods'. This work is supported by the Bank of Sweden Tercentenary Foundation.

REFERENCES

Bruce, G. 1982. 'Developing the Swedish intonation model'. *Working Papers* 22, 51-116. Department of Linguistics, Lund University

Engen, T., E. Engen, R. Clarkson, & P. Blackwell. 1983. 'Discrimination of intonation by hearing impaired children'. *Applied Psycholingusitics* 4, 149-160.

Fónagy, I. 1981. 'Emotions, Voice and Music'. *Research Aspects of Singing* 33, 51-79. Swedish Academy of Music

Fourcin, A.J. 1980. 'Speech pattern audiometry'. *Auditory investigation; the Scientific and Technological Basis,* ed. H. A. Beagley,170-208. Oxford: Clarendon Press.

Gårding, E. 1986. 'Intonation och Sinnestämningar'. Föredrag vid Humanistdagarna 1986. Lund University.

Gårding, E. & A. Abramson. 1965. 'A study of the perception of some American English intonation contours'. *Studia Linguistica* 19, 61-79.

Hadding - Koch, K. 1961. *Acoustico-phonetic studies in the intonation of Southern Swedish.* Lund: Gleerup

House, D. 1990. 'On the perception of mood in speech'. *Phonum* 1, 113-116. Department of Phonetics, University of Umeå.

House, D., G. Bruce, L. Eriksson, & F. Lacerda 1988. 'Recognition of Prosodic Categories in Swedish: Rule Implementation'. *Working Papers* 33, 153-161. Department of Linguistics and Phonetics, Lund University.

Öster, A-M., & A. Risberg. 1986. 'The identification of the mood of a speaker by hearing impaired listeners'. *Speech Transmission Laboratory - Quarterly Progress and Status Report (STL-QPSR)* 4/1986, 79-90. Department of Speech Communication and Music Acoustics, Royal Institute of Technology, Stockholm.

Risberg, A. & E. Agelfors. 1978. 'On the identification of intonation contours by hearing impaired listeners'. *Speech Transmission Laboratory - Quarterly Progress and Status Report (STL-QPSR)* 2-3/1978; 51-61. Department of Speech Communication and Music Acoustics, Royal Institute of Technology, Stockholm.

Waltzman, S. & I. Hochberg. 1990. 'Perception of Speech Pattern Contrasts Using a Multichannel Cochlear Implant'. *Ear and Hearing* 11:1, 50-55.

Williams, C.E. & K.N. Stevens. 1972. 'Emotions and speech: Some acoustical correlates'. *The Journal of the Acoustical Society of America* 52:4, 1238 -1250

Lund University, Dept of Linguistics
Working Papers 36 (1990), 109-121

Once a Poor Reader - Always a Poor Reader?

Eva Magnusson and Kerstin Nauclér

This paper was presented at the 2nd meeting of the European Group for Child Language Disorders on March 17 - 21, 1990 in Röros, Norway. The research is supported by a grant from HSFR (Swedish Council for Research in the Humanities and Social Sciences).

INTRODUCTION

In this paper we will present some data on the development of both language-disordered and normally speaking children's reading abilities from grade 1 to grade 3. The ultimate aim of the research project is to predict reading and spelling achievements and to develop methods for identifying at-risk children at an early age, preferably before they learn to read and write. If the children can be found already as preschoolers, they can be given adequate training before they start school instead of having to live through the slow and painful experience of winning a reputation as "poor readers", "bad students" and "unable to learn to read and write" before they finally qualify for remedial teaching.

Before going into the questions to be discussed in this paper, we will give some background information about the project. The children's reading and spelling achievements in grades 1, 3 and 4 are correlated with data from their preschool years on linguistic ability, linguistic awareness and other abilities that are considered important for the development of reading and writing such as short-term memory and information processing strategies.

When we started six years ago the subjects were 115 children, 76 language-disordered children and 39 normally speaking children. The language-disordered children consisted of all the children in the city of Malmö, Sweden, born in 1978, who were diagnosed as having an idiopathic (functional) language disorder. The language-disordered children fell into two subgroups, one which had language therapy before starting school (39 children) and one group (37 children) which did not, depending on whether the disorder was diagnosed by a speech pathologist as severe or mild. The 39 children in the normal group had been matched on an individual basis to the children with the most severe language disorders, i.e. the language-