

TACTILE TRANSMISSION OF INTONATION AND STRESS

Hans Georg Piroth and Thomas Arnhold
Institut für Phonetik und Sprachliche Kommunikation
der Universität München

Schellingstr. 3, D-80799 Munich, Germany
Tel. +49 89 2178 2655, Fax +49 89 2178 2652, E-mail: piroth@phonetik.uni-muenchen.de

ABSTRACT

The development of technical communication devices for hearing-impaired or deaf persons is one of the main topics in actual research on aids for the disabled and elderly. Besides the well-known advances in the construction of analogue and digital hearing aids for the hard-of-hearing, there is also a long tradition of research on tactile speech aids. From the beginning of this century, many attempts have been made to resolve the problem of speech substitution for deaf people and those suffering from severe hearing loss. A series of experiments was carried out to determine a robust coding method for tactile transmission of F₀ and stress to support speech perception of deaf or severely hearing impaired persons unable to extract suprasegmental speech features by the auditory sense.

1. INTRODUCTION

The work presented here is part of the results of a European Union research project (Telematics for the Disabled and Elderly - TIDE). Within the framework of this project (Signal Conditioning Communication Aids for the Hearing Impaired - SICONA) [1] it is argued that the selection of the tactile senses as a channel for transmitting suprasegmental features would facilitate the use of residual hearing for the auditory perception of segmental cues. It is the concern of this article to determine a robust electro-tactile coding method for the cutaneous transmission of intonation (F₀ variation) and rhythm (loudness variation) of speech.

Experiments carried out in the thirties of this century (*e. g.* Monjé [2]) have already shown that frequency differences in cutaneous stimulation are perceivable when a.c.-frequencies below 200 Hz are used.

Classical investigations on electro-tactile intensity perception have also discussed the number of possible steps that can be discriminated. Lindner [3] has reported that the pain threshold is reached at approximately four times absolute threshold. At a frequency of 400 Hz he situates absolute threshold at about 0.8 mA, pain threshold at 4.7 mA with 27 discriminable steps inbetween. Schöbel [4] determined a difference limen of 4 to 5%, Anderson and Munson [5] of 2 to 5% in the

frequency range between 100 and 5000 Hz. Hawkes [6] measured a limen of 5.3% at an intensity of 120% and of 3.8% at 200% above absolute threshold.

Nevertheless, in order to function as a coding feature for electrocutaneous speech transmission in communication aids to be used in daily life, mere discriminability of intensity steps in standard discrimination tests is not sufficient. When intensity is a feature of a complex, speech coding system variations have to be clearly identifiable even in a noisy background. Earlier investigations of one of the authors [7] have already shown that identifiability of only two intensity levels is stable enough to encourage the implementation of such a feature in complex multimodal speech patterns.

2. EXPERIMENTS

2.1 Apparatus and Stimuli

The experiments were carried out with the electrocutaneous stimulation device SEHR-3 [8] that produces bipolar rectangular impulses of variable amplitudes and durations freely arrangeable in sequences. Impulses are delivered to the skin by pairs of circular gold-layered (5 μ) brass electrodes (9 mm in diameter each with a center-to-center distance of 10 mm between the electrodes of each pair).

According to pretest results impulse durations were set to 208 μ s. In advance of each testrun absolute (T_{AB}) and annoyance (T_{AN}) thresholds were determined for each participant. Before stimulus presentation, amplitude was preset individually to the mid value between thresholds ($I_{MID} = T_{AN} - T_{AB}$). Again, based on earlier investigations and for practical reasons, the dorsal side of the left forearm was chosen as stimulation site [9].

Whenever an attempt is made to present selected features of the speech signal via an alternative sensory modality (suprasegmental features of intonation and stress in the case of the investigation described here) it must be ensured that the user of such an aid is able to integrate the multimodal sensory input during the processing of the coded speech. The coding method proposed tries to enhance this ability by preserving natural speech cues as far as possible. Though using the tactile sense as an alternative modality for speech perception variations of

natural speech are transposed into variations of the same stimulus dimensions of the alternative sense. *I. e.* to tactually present intonation and stress, acoustic fundamental frequency is mapped onto tactile impulse repetition rate and dynamic accent onto tactile intensity.

To yield comfortable and well-discriminable tactile patterns for the transmission of F0 information impulse sequences oscillating between two electrode pairs 100 mm apart from one another were used for the experiments. These stimuli were presented with pulse repetition rates (pps) that create the sensation of an apparent sweeping movement between these electrode pairs.

2.2 Tactile Coding of Intonation (Experiment 1)

Experiment 1 was designed to test the identifiability of final frequency rises and falls in tactile impulse sequences. The baseline frequency was set to 50 pps corresponding to half average male speech F0, remained constant during the first 350 ms of the stimulus and was raised or lowered over 250 ms to the end of the sequence. An experiment previously reported [10] showed that in such stimuli rises and falls to the extent of 20 pps resulted in identification rates of less than 70%.

Based on these results, a 60 pps range of frequency variation was used in the present experiment (baseline 50 Hz, target frequencies in the final part of the stimuli 90, 80, 70, 60, 50, 45, 40, 35, and 30 Hz). 8 repetitions of these 9 stimuli were presented in random order with an interstimulus interval of 5 s. 7 participants underwent the test procedure and were asked to identify whether the stimuli ended in a rise or a fall.

Fig. 1 shows identification rates of more than 80% for the outmost rises and falls.

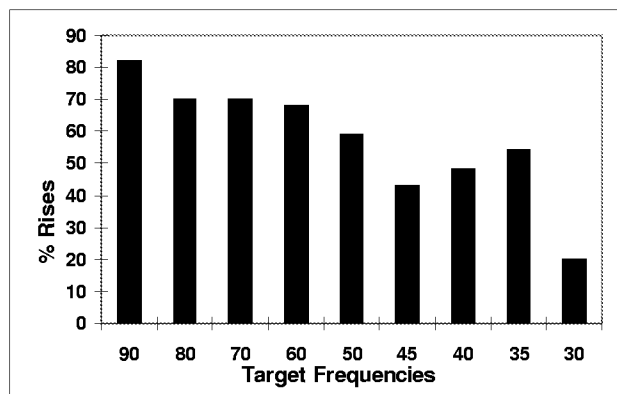


Figure 1: Identification of Final Frequency Rises and Falls in Experiment 1 (% Identified Rises)

2.3 Tactile Coding of Stress (Experiment 2)

To code a tactile equivalent of stress an additional variation of intensity was superimposed on the initial constant frequency part of the stimuli of Experiment 1.

8 repetitions of 27 stimuli were arranged in random order for Experiment 2: In one third of the stimuli the mid intensity level calculated ($I_{MID} = T_{AN} - T_{AB}$) was preserved throughout the impulse sequences. For the other two thirds intensity peaks were implemented with an amplitude maximum just below the upper threshold for one half of the stimuli ($I_{PEAK} = I_{MID} + 0.40 (T_{AN} - T_{AB})$) and an intensity increase of about 50% of this amount for the other half ($I_{PEAK} = I_{MID} + 0.25 (T_{AN} - T_{AB})$).

In Experiment 2a 6 participants were asked to identify the presence or absence of intensity peaks. Figure 2 gives the percentages of identified peaks for the three stimulus conditions. It can be seen that the recognition of intensity variations in these stimuli is rather poor and reaches nearly 70% if peak-intensity is raised to near the threshold of annoyance value.

It can be suggested that simple insertion of intensity peaks is insufficient to mark word accents tactually for practical use.

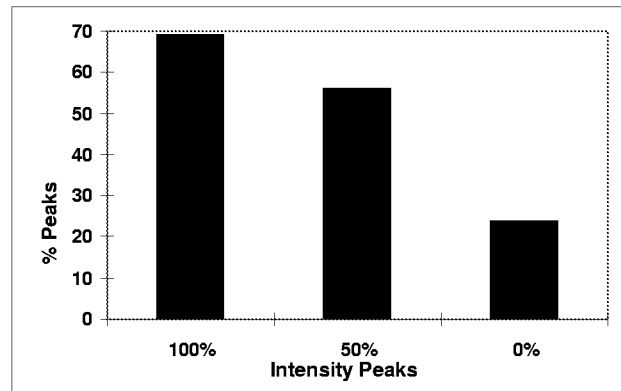


Figure 2: Identification of Initial Intensity Peaks in Experiment 2a (% Identified Peaks)

In a control experiment (Experiment 2b) using the same stimuli and procedure 6 participants had to identify whether the tactile patterns presented ended in a rise or fall. Fig. 3 shows that the recognition rates for frequency variations in the final part of the stimuli resembles to the results of Experiment 1.

Thus, it can be concluded that the co-occurrence of frequency and intensity variations does not severely disturb the identifiability of the encoded frequency variations.

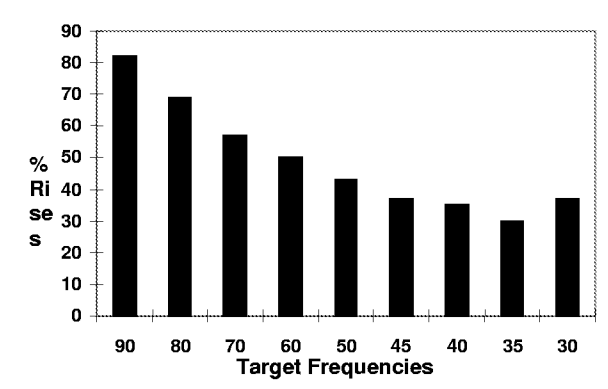


Figure 3: Identification of Final Frequency Rises and Falls in Experiment 2b (% Identified Rises)

Experiment 2c was designed to determine whether identification of intensity peaks improves, if no additional suprasegmental variation is present in the stimulus patterns.

Impulse sequences with a constant frequency of 50 pps and an overall duration of 680 ms were constructed with intensity characteristics as described for Experiment 2a. Again, Fig. 4 presents the rates for the identification of intensity peaks (8 subjects).

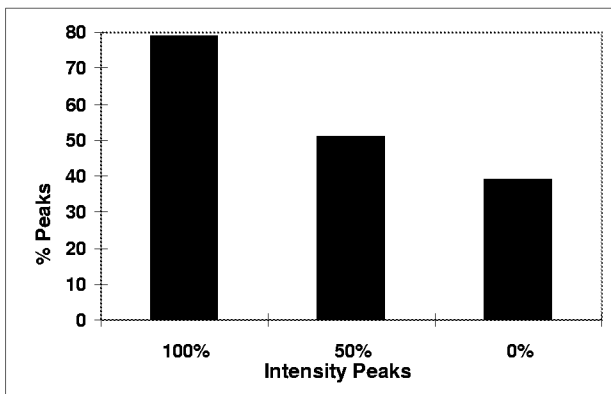


Figure 4: Identification of Initial Intensity Peaks in Experiment 2c (% Identified Peaks)

As can be seen from the diagram, recognition rate is higher than in Experiment 2a, but does not exceed 80% even for the most prominent peak. Also the large number of false alarms (peaks detected in completely level stimuli) is noteworthy.

Summarizing the results of Experiments 1 and 2 it can be stated that both tactile coding dimensions - frequency and intensity - can be recognized at an 80%-level when presented alone for the clearest stimulus conditions. Identification rates show only a slight decay when variations of both parameters co-occur. Nevertheless, identifiability seems to be low with regard to the

demands for practical applicability in realtime communication situations or in noisy environments.

2.4 Concomittant Variations of Tactile Frequency and Intensity (Experiment 3)

To yield robust identification rates sufficient for practical use in non-laboratory surroundings the combination of both stimulus parameters to a complex coding feature has been investigated.

In the execution of Experiment 3 the same stimulus set was used as in Experiment 1, but with simultaneously increasing or decreasing intensities. For the highest F0 rise intensity concomittantly increased to a value approximately 10% below the threshold of annoyance ($I_{TARGET} = I_{MID} + 0.40 (T_{AN} - T_{AB})$), for the deepest fall it decreased to 10% above absolute threshold ($I_{TARGET} = I_{MID} - 0.40 (T_{AN} - T_{AB})$). F0-variations of the other stimuli within this range were accompanied by intensity variations to the appropriate target values inbetween. Number of participants and procedure were the same as in Experiment 1.

As can be seen from Fig. 5, the intensity-enhanced F0 rises and falls could be clearly identified with rates reaching nearly 100% for the edge stimuli with the largest amounts of variation.

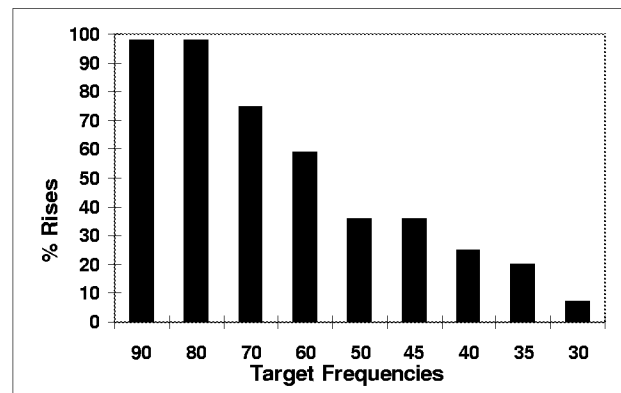


Figure 5: Identification of Final Rises and Falls in Experiment 3 (% Identified Rises)

3. CONCLUSION

The investigation revealed that the combined variation of (normalized) electro-tactile pulse repetition rate and (individually adjusted) amplitude of electric impulses yields well-recognizable patterns for speech-to-skin transmission of suprasegmentals. This coding method can be realized without intensive technical effort and can easily be explained to users. Nevertheless, it must be emphasized that only the combination of both "naturally" transposable features (frequency and intensity) is robust enough to ensure the applicability under everyday life conditions. Thus, the final decision has still to be made

how suprasegmental features of speech could be tactually presented. One should consider transmitting both features - intonation and rhythm - by the same tactile method - combined variation of tactile impulse frequency and intensity - since intensity and F0 co-vary also in spoken language to a great extent. However, it has to be taken into account that perception of tactile frequency variations alone and tactile intensity variations alone was poorer, but at least recognition rates reached up to 80%. Thus, under special conditions (*e. g.* in a quiet environment or if the user does not move during tactually aided communication) a separate coding of intonation by tactile pulse repetition rate and of rhythm by tactile intensity might be sufficient for the purpose of enhancing speech perception by hearing-impaired persons.

4. ACKNOWLEDGMENT

Work was supported by the Commission of the European Communities (TIDE-Project No. 1090: SICONA).

REFERENCES

- [1] D. Bauer, J. C. Geiger and R. Beerwerth, "Speech Signal Conditioning Communication Aids for the Impaired with Severe Auditory Sensory Damages", in: "Speech and Language Technology for Disabled Persons", ed. by B. Granström, Sh. Hunnicutt & K. E. Spens, pp. 51-54, Stockholm (ESCA), 1993.
- [2] M. Monjé, "Über die Wirkung von Wechselströmen verschiedener Frequenz auf die Hautsensibilität", *Zeitschrift für Sinnesphysiologie*, vol. 67, pp. 2-18, 1936.
- [3] R. Lindner, "Physiologische Grundlagen zum elektrischen Sprachetasten und ihre Anwendung auf den Taubstummenunterricht", *Zeitschrift für Sinnesphysiologie*, vol. 67, 1937.
- [4] E. Schöbel, "Versuche über Intensitätsunterscheidung beim 'elektrischen Tasten' verschiedener Frequenzen", *Zeitschrift für Sinnesphysiologie*, vol. 66, pp. 262-273, 1936.
- [5] A. B. Anderson and W. A. Munson, "Electrical Excitation of Nerves in the Skin at Audiofrequencies", *Journal of the Acoustical Society of America* vol. 23, p. 155, 1951.
- [6] G. R. Hawkes, "Cutaneous Discrimination of Electrical Intensity", Phil. Diss. Univ. of Virginia, 1959.
- [7] H. G. Piroth, "Intensity Discrimination and Identification in Electrotactile Pulse Train Sequences", Forschungsberichte des Instituts für Phonetik und Sprachliche Kommunikation der Universität München, vol.30, pp. 65-78, München, 1992.
- [8] H. G. Piroth and H. G. Tillmann, "Das System zur elektrischen Hautreizung SEHR-3", Forschungsberichte des Instituts für Phonetik und Sprachliche Kommunikation der Universität München, vol.29, pp. 41-46, München, 1991.
- [9] H. G. Piroth, "Elektrotaktile Silbenerkennung mit quasi-artikulatorisch kodierten komplexen zeitlich-räumlich strukturierten Reizmustern", Forschungsberichte des Instituts für Phonetik und Sprachliche Kommunikation der Universität München, vol. 22 (Phil. Diss.), München, 1985.
- [10] H. G. Piroth and T. Arnhold, "A Supplementary Tactile Speech Aid Transmitting F0-Information", Proc. 13th ICPhS, vol. 4, pp.516-519, Stockholm, 1995.