POSSIBILITIES AND LIMITS OF USING SPEECH SIGNALS IN AVIATION AND SPACE PSYCHOPHYSIOLOGY

J. ŠULC and V. REMEK
Institute of Aviation Medicine
Kovpakova 1, 160 60 Prague 6, Czechoslovakia

Key words: emotion, space psychophysiology, speech

Abstract. A number of measurable speech characteristics reflect the emotional state and its influence upon the operator's performance. The validity of acoustical and temporal measures of paralinguistic and semantic processes increases with the increase of emotional activation. At lower activation levels the informational value of characteristics is mostly affected by linguistic factors as well as by psychological ones, related to social communication in the course of common working activity.

Human speech is capable of transferring a remarkably rich set of complex and highly integrated information on various physical and psychological characteristics of the speakers. In the course of communication the human brain actively classifies and extracts information relevant to the needs and interests of the acting subject. In comparison with other instrumental methods used in psychophysiological research, speech signal processing has certain advantages. There is no need of special sensors, the examination does not make the subject feel that he is observed. The acquisition as well as the processing and storing of the records is very simple.

Already two decades ago these benefits led the Soviet specialists in aviation and space medicine to monitoring some psychological variables of the pilot or cosmonaut in the course of flight by means of speech sig-
nals (1, 7). In the following years the main interest of the investigators was directed to the determination of acoustical correlates of emotional speech in various dangerous in-flight situations (2, 6, 8, 12, 16, 18, 19). Due to the complexity of encoding the emotions and other psychological variables into particular speech parameters, appropriate attention has been devoted also to other psycholinguistic measures (9, 10, 11). Our experience in this field during the last 15 years will be discussed in this paper.

The utterances of the aircrew, spacecrew and ground personnel were obtained from routine tape-recordings of ongoing radiophony. They were processed by means of a Sona-Graph 6061 B Spectrum Analyzer, using the method described by Hecker et al. (4). Some tape-recordings have been processed by means of a Recorded Signal Detector (3) and a LINC 8 DEC laboratory computer. In a portion of the acoustic material a content analysis of utterances was performed according to rules explained in detail elsewhere (10, 11). A wide set of complementary information on the crew's behavior facilitated the appropriate interpretation of the results of acoustic and psycholinguistic analysis.

Due to the specific nature of the acoustical signal transmitted from the aircraft or the spaceship (clipped upper frequencies in spectrum, considerable admixture of noise and echoes from cockpit or cabin surfaces etc.), only the following acoustic and temporal parameters were measured: the pitch $F_0$ (4), the band width of voice $D$ and the output of speech per breath $ER$ (3). Four groups of semantic classes were distinguished in the cooperation between the crew and ground operator, viz. the directive, informative, interrogative and deliberative utterance (10, 11).

The nature of acoustic material was inevitably very heterogenous. From the point of view of the concrete situation in the context of the whole material, it could be divided into recordings related to emergencies and those related to general operations. Conclusions concerning acoustical and psycholinguistic peculiarities of speech, produced in situations of imminent threat to one's own or somebody else's life, were based on a detailed analysis of 100 real in-flight incident/accident situations. Our analysis is based upon psychoacoustical features of standard aviation and cosmic radiophony obtained from whole military solo flights and of the Soyuz 28-Salyut 6 mission, representing 8133 utterances.

It has been confirmed that the most important information about emotions is mediated through melodic features of the speech signal (3, 5, 13, 19). The pitch always increased significantly already in the first utterance, which followed the onset of an unexpected emergency situation. The increase usually reached 40-50% above the baseline. The voice
band width behaved similarly, it increased by 50-80% above the baseline.

ER during speech varied in a more complicated manner. In subjects solving the situation "in cold blood" it increased regularly. On the contrary, psychoemotional inhibition and passivity with respect to the situation manifested itself in the slowing down of speech. The deviations in a positive or negative direction reached about 15-30% (13, 14).

Changes of speech related to the actual emotional state and the operator's working efficiency can be demonstrated in the following example (Fig. 1). A young, inexperienced jet fighter pilot performed a series of serial attacks on a ground target. All 5 attempts were unsuccessful and during the last one of the two engines stopped at a low altitude. The emergency situation, until the pilot's ejection, lasted for 36 s. During the first 30 s the pilot presented signs of strong psychic and emotional inhibition. He was helpless and resigned to passive execution of instructions, which he got from the flying control officer. During all main phases of the flight emotional activation was always connected with changes in the pitch and ER. In this case asthenic emotions with gradual block of performance were combined with a characteristic dissociation between increasing $F_0$ and relatively low ER. Only in the last 6 s, after the decision of leaving the aircraft by ejection, his voice reflected, along with the outlasting negative emotion, a reversal from passive to active action.

![Fig. 1. Changes of speech related to actual emotional state. $F_0$, expressed in Hz (white circles), and ER, (black circles) accompanying shame, fear and improper performance. ER, the ratio of the number of syllables to the time of their expulsiv per one expiration. A, launch; B-F, unsuccessful ground target attacks; G, engine stoppage; H, pilot ejection.](image)
Besides vocal indicators of psychological stress, there are other important signs of emotional activation (e.g. hesitation patterns, slips of the tongue, syntactic alterations, etc.). They assert themselves even in such a formalized and regularized verbal communication system as the aviation and cosmic radiophony. However, for the time being the informational value of these characteristics does not reach the importance of acoustical and temporal parameters discussed.

In this connection these arises a question of the margins between non-specific and emotion-signalling features in a speech signal. Recently exploratory studies have been performed to investigate this topic (15, 17). From an analysis of 2594 utterances delivered during 6 h of uncomplicated local flights by 29 pilots and 1 flight control officer it followed that the value of ER in non-emotive dyadic communications depends significantly on the semantic component of the utterances (Fig. 2). Informative utterances were produced at a slower rate ($P < 0.01$, t-test) than interrogative ones. This finding indicates that variability in ER is affected, besides the dynamics of conscious regulation of behavior, also by factors related to the type of social interaction among speakers engaged in a common activity.

This hypothesis has been confirmed by the analysis of a co-operative dialogue between the Soyuz 28-Salyut 6 spaceship crew and the operator of mission control centre (17). The pitch as well as the ER were significantly lower ($F_0$ by $15-24\%$, $P < 0.02$ and ER by $27\%$, $P < 0.05$, t-test) in the spacecraft commander's utterance, directed to the co-pilot in contrast to identical communications broadcasted to the ground operator (Fig. 3). This tendency has not been found in the co-pilot. The greater flexibility of extralinguistic characteristics in the speech of spacecraft commander points to the differences in his actual social roles.
Fig. 3. Dependence of $F_o$ during informative utterances on sociometric relations, A, placing in orbit; B and C, docking maneuver; D, re-entry. SCDR, spaceship commander; OSC, oscilloscope spaceship co-pilot. White bars, space-craft-to-earth communications; stripped bars, on-board face-to-face communications.

with respect to his coworkers. However, more complete studies would be necessary to differentiate between speech characteristics typical for normal conditions and those typical for activation levels, considered yet as debilitating ones.'


Accepted 20 June 1988