Effects of Noise on Identification of Topic Changes in Discourse

G.A. Tidball, M.K. Pichora-Fuller, J.H.V. Gilbert, N.Lamb
School of Audiology and Speech Sciences
University of British Columbia
Vancouver, British Columbia, Canada

RATIONALE

Hard-of-hearing individuals frequently report that they are able to understand speech when the topic is known but that they experience difficulty when the topic changes. The motivation for the present study was to investigate how topic identification might influence listening comprehension in unfavourable listening conditions.

Listeners rely on contextual information in speech to aid comprehension. The richer the context, the less a listener must rely on the perception of the acoustic features of speech to comprehend what is heard.

An example of contextual knowledge which aids comprehension is knowledge of topic, or, simply stated, "what is being talked about." Previous studies have shown that comprehension and retention of language are aided by individuals' prior knowledge of topic (Bransford & Johnson, 1973; Larch & Larch, 1985). Knowledge of topic provides constraints on the language content of discourse. It may also help individuals to organize discourse content according to some pre-existing concept of language structure such as a hierarchical structure of topic and sub-topics.

While topic may be defined with respect to content, it may also be structurally defined. That is, in addition to knowing what the topic is, listeners are also aware of when a new topic is beginning or an old topic is continuing or ending. The boundaries of topics in discourse are indicated lexically, syntactically, and perhaps most importantly prosodically (e.g., variations in voice pitch, loudness, pausing, etc.). It follows, then, that in order for listeners to be able to identify a topic of discourse, they must be able to identify the boundaries of discourse topic, or when "what is being talked about" has changed. Inability to correctly perceive those structural cues which signal topic transitions would deprive the listener of important contextual information from which he/she could predict the content of the discourse.

The purpose of the present study was first, to determine how accurately normal-hearing individuals identify where topic changes occur in discourse under favourable listening conditions; and second, to determine how normal-hearing individuals' ability to identify topic changes in discourse is affected by competing noise.

METHODS

The stimulus materials used in the study were recordings of monologues spoken by a young, native English-speaking adult male. Monologues were employed to elicit natural discourse in which there was no turn-taking which might cue subjects to a change in topic.

The talker was asked to describe seventy-two photos chosen by both the experimenter and the talker from a large selection of the talker's family photos. Each picture was of a different event. The talker was instructed to describe the pictures so that future listeners might be able to imagine for themselves what each picture looked like. He was also asked to describe each photograph individually, without referring to any information mentioned in the descriptions of previous photos. Photographs were viewed one at a time. For the purpose of this study, each photograph was considered to represent a topic.

The talker was located in a sound-attenuating booth during the elicitation of the stimulus materials. The monologues were recorded digitally using SoundWorks (v.3) on a NeXT computer system.

Twelve normal-hearing listeners participated in the perceptual study. All were native Canadian English speakers between the ages of 21 and 35 years.

Prior to the test session, subjects were familiarized with the talker's voice by listening to a recording of him speaking on a topic unrelated to the test materials. Subjects then listened to the stimulus monologues in three signal-to-noise (S:N) conditions: +5 dB S:N (favourable listening condition), 0 dB S:N (less favourable listening condition), and a -5 dB S:N (unfavourable listening condition). Each monologue was comprised of eleven photograph descriptions, yielding a total of ten topic changes per condition. Half the subjects listened to the unfavourable listening condition (-5 dB S:N) first, half listened to the most favourable listening condition (+5 dB S:N) first. The monologues were presented monaurally to the subject's better ear.

The subjects' task was to indicate when they believed the talker in the pre-recorded monologue was about to begin talking about a new photograph. Subjects' indicated their responses by pressing a button placed on a table in front of them. This generated a response signal. The response signal together with the stimulus materials were recorded simultaneously onto an analogue tape cassette.

ANALYSIS

The analogue recordings of the stimulus materials and subjects' responses were re-recorded onto the NeXT sound processing system for analysis. We measured the latency of the subjects' responses in relation to the conclusion of one topic or photo description and the beginning of the following topic.

Three primary measures were taken:

1. the median latency of each subject's responses in each condition;
2. the number of false positive identifications of topic changes;
3. the number of times subjects waited until the following photo description or topic had begun before they indicated that they recognized a topic change.

An increase in one or all of these measures could indicate increased difficulty in subjects' ability to perform the experimental task. We measured the variation of subjects' ability to perform the experimental task as a function of increased competing noise.

RESULTS

Significant differences in subjects' median latency of response were found between the +5 dB and 0 dB S:N conditions ($p < .05$) and between the +5 dB and -5 dB S:N conditions ($p < .01$). No significant difference in median latency was found between the 0 dB and -5 dB S:N conditions ($p > .05$).

A significantly greater number of false positive responses were given in the -5 dB S:N condition than were given in either the 0 dB or the +5 dB S:N conditions ($p < .0001$).

The number of times subjects waited until the following photo description or topic had begun before they indicated that they recognized a topic change was also found to vary as a function of signal-to-noise condition. A significantly greater number of this type of response was given in the -5 dB S:N condition than were given in either the 0 dB and +5 dB S:N conditions ($p < .01$).

DISCUSSION

The above results show that in a favourable listening condition (+5 dB S:N), listeners are quite accurate in anticipating topic changes in discourse with some degree of accuracy. As the listening condition becomes increasingly unfavourable, listeners (1) require more time to identify the boundaries of topics, (2) rely more heavily on those cues which indicate initiation than those which indicate termination of a topic, (3) make more errors in their estimates of where topic boundaries occur.

We offer the following explanations as to why these results were obtained: that reduction of segmental and suprasegmental information in the unfavourable listening conditions reduces listeners' ability to perceive those cues necessary to identify topic changes in discourse; that decreased cues to topic boundaries results in increased processing time - the listener must listen longer to obtain enough information to identify topic boundaries; and that increased competing noise may mask parts of the speech signal to such a degree that the listener no longer detects the signal, with an apparent pause resulting in "miscues".

These results suggest that listeners have difficulty identifying the cues necessary to identify topic boundaries in discourse when the speech signal is significantly degraded by competing noise. Difficulty identifying topic boundaries decreases listeners' ability to organize discourse into topics, and thereby may reduce contextual information in speech.

REFERENCES

Bransford, J.D., & Johnson, M.K. (1973). Considerations of some problems of comprehension. In W.G. Chase (Ed.), *Visual Information Processing*. New York: Academic Press.

Larch, R.F., & Larch, E.P. (1985). Topic structure representation and text recall. *Journal of Educational Psychology*, 77, 137-148.

Figure 1 Mean of the Median Latency of Subjects' Responses

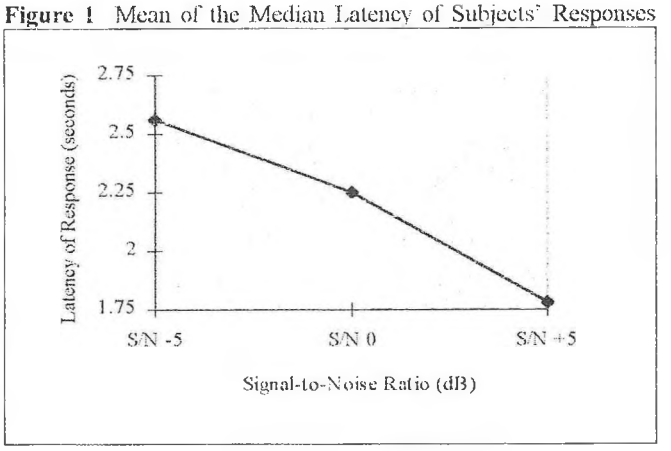

Figure 2 Mean Number of False Positive Responses

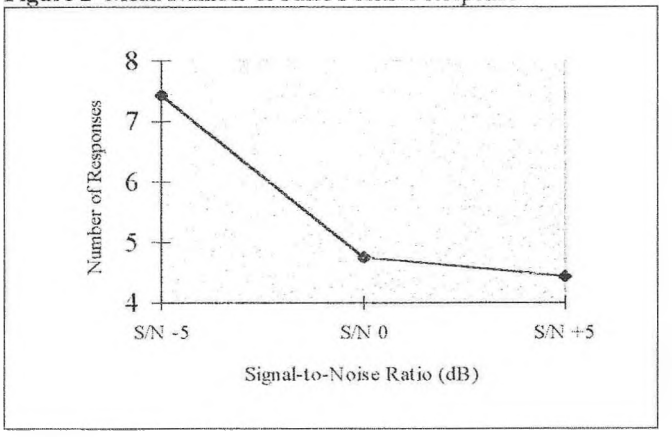

Figure 3 Mean Number of Times Subjects Waited Until Following Topic Began Before Responding