

RESEARCH ON SPOKEN LANGUAGE PROCESSING
Progress Report No. 24 (2000)
Indiana University

Change Deafness: The Inability to Detect Changes in a Talker's Voice¹

Michael S. Vitevitch

*Speech Research Laboratory
Department of Psychology
Indiana University
Bloomington, Indiana 47405*

¹ This research was supported, in part, by NIH-NIDCD Training Grant T32 DC 00012 to Indiana University. I would like to thank Robert Nosofsky, David Pisoni and Holly Storkel for helpful comments and suggestions.

Change Deafness: The Inability to Detect Changes in a Talker's Voice

Abstract. Change blindness is a failure to detect a change in a visual scene. A shadowing task was used to demonstrate an auditory analogue to change blindness—*change deafness*. Participants repeated words varying in lexical difficulty. After a rest-break they heard more words from either the same or a different talker. Answers to explicit questions about the change in talker and implicit measures of behavior (i.e., response latencies) demonstrate that processing is affected by the change, even if participants do not explicitly report a change in talker. Specifically, listeners who did not detect the change in the talker had a greater difference between conditions of lexical difficulty than listeners who noticed the change, or listeners who heard the same talker throughout. These results suggest that failures to detect changes are not limited to the visual domain and that processing at some level may be affected by changes in the environment.

Change blindness is a counterintuitive phenomenon (Levin, Momen, Drivdahl & Simons, in press; Simons & Levin, 1998) in which observers in a variety of paradigms (e.g., Henderson & Hollingworth, 1999; Levin & Simon, 1997; Rensink, O'Regan & Clark, 1997) fail to detect what may be described as obvious changes in the environment. For example, Grimes (1996) found that participants noticed only 30% of the changes in photographs that occurred during an eye movement—even changes as obvious as two heads switching bodies. Simons and Levin (1998) dramatically demonstrated that only 33% of the participants in a real-life interaction noticed that the person asking them for directions was exchanged when a door being carried by confederates momentarily interrupted the discussion.

Is the inability to detect changes in the environment unique to the intense processing demands of the visual system that must encode complex visual-spatial details during very brief eye-fixations, or are there analogous deficits to detecting changes in similarly complex “scenes” in other modalities? In the auditory domain, spoken language may be a comparably complex stimulus. Speech is a complicated auditory signal that simultaneously conveys a conceptual, linguistic message and indexical information to a listener. Indexical information refers to acoustic correlates in the speech signal that provide information on various characteristics of the talker including identity, emotional state, age, dialect, and gender (Pisoni, 1997).

A number of studies have found that changes in the talker producing the stimulus words affects the accuracy with which participants identify those words presented in noise (Mullennix, Pisoni & Martin, 1989; Nygaard, Sommers & Pisoni, 1994), as well as the later recall and recognition of stimulus words (Martin, Mullennix, Pisoni & Sommers, 1989; Palmeri, Goldinger & Pisoni, 1993; Goldinger, Pisoni & Logan, 1991). Furthermore, Goldinger (1998) has shown that words spoken by the same talker in training and test sessions are repeated faster than words spoken by different talkers between training and testing sessions. The results of these studies suggest that changes in the voice of the talker affect processing. Unfortunately, in none of these experiments were participants *explicitly* interrogated to see if they detected the change in the talker that produced the stimuli. Thus, it is unknown if participants explicitly noticed the change in talker, or if they were “deaf” to this change. The results from change blindness studies (e.g., Simons & Levin; 1998) might lead one to predict that most of the participants in the talker-variability studies were unaware of the change in talkers. In contrast, the results from the talker-variability studies clearly demonstrate that participants' responses were affected by the change in the talkers (e.g., Palmeri, Goldinger & Pisoni, 1993).

To reconcile this apparent contradiction the present experiment measured the response latencies of participants to repeat words that were produced by the same talker throughout the experiment or by talkers that were changed halfway through the experiment. Furthermore, participants were *explicitly* asked at the end of the experiment if they noticed the change in the talker. By using both implicit and explicit measures, the present experiment addresses important questions regarding change detection and talker-variability. Specifically, is the inability to detect changes limited to the visual domain, or do participants also exhibit change deafness to changes in the auditory environment? More interestingly, this experiment will allow us to see if behavior is affected *implicitly* by the change in the talker (via differences in reaction times across groups) even if participants do not *explicitly* detect the change (see Chun & Nakayama, in press; Hayhoe, in press; Williams & Simons, in press).

Method

Participants

Twenty-four native speakers of English who reported no history of hearing or speech disorders participated in the experiment for partial fulfillment of an Introductory Psychology research requirement.

Stimuli

One hundred words with a familiarity rating of 6 or higher on a seven-point scale (Nusbaum, Pisoni & Davis, 1984) were selected for this experiment from the Indiana “Easy-Hard” Multi-Talker Speech Database (Torretta, 1995). Fifty words were lexically *easy* and fifty words were lexically *hard*. Lexically easy words have high word frequency and few similar sounding words with a low frequency of occurrence, whereas lexically hard words have low word frequency and many similar sounding words with a high frequency of occurrence (Torretta, 1995). These variables were statistically different in this subset of stimuli. The mean word frequency (based on word counts from Kucera & Francis, 1967) for easy words was 173.02 occurrences per million, and 8.5 occurrences per million for hard words ($F(1, 96) = 41.75, p < .001$). The mean number of similar sounding words, or neighbors, for easy words was 13.36 neighbors, and 27.24 neighbors for hard words ($F(1, 96) = 238.71, p < .001$). The mean frequency of the neighbors for easy words was 34.68 occurrences per million, and 302.19 occurrences per million for hard words ($F(1, 96) = 58.42, p < .001$). The results from a number of different behavioral tasks and participant populations show that, in general, participants respond more quickly and more accurately to lexically easy words than to lexically hard words (e.g., Kirk, Pisoni, Miyamoto, 1997; Luce & Pisoni, 1998; Sommers, 1996).

The same one hundred words were selected from two different male talkers in the database (talkers M0 and M9). These words were pre-tested by ten additional listeners from the same population in an AX “same-different” task to confirm that the selected talkers were perceptually discriminable. These participants heard the same word twice (separated by 50 ms. of silence). The word was spoken either by the same talker or by the two different talkers. When the word was produced by the same talker, participants were 98.6% accurate in responding that the voices were the same. When the word was produced by the two different talkers, participants were 92.2% accurate in responding that the voices were different. These results suggest that the two male voices were highly discriminable perceptually and that any failures to detect the change in the voice were not due to the perceptual similarity of the voices of the talkers.

Procedure

Participants were tested one at a time on a Macintosh Quadra 950 running PsyScope 1.2.2 (Cohen, MacWhinney, Flatt & Provost, 1993) which controlled stimulus randomization and presentation, and collection of response latencies. A headphone-mounted microphone (Beyer-Dynamic DT109) was interfaced to a PsyScope button box that acted as a voice-key. A typical trial proceeded as follows: A stimulus word was presented over the headphones to a participant who had been instructed to repeat the word as quickly and as accurately as possible. Response latency, measured from the beginning of the stimulus, was triggered by the onset of the participant's verbal response. Another trial began 1 s after a response was made. Responses were also recorded on audio-tape for later accuracy analyses.

Each participant received a total of 100 trials. In the first half of the experiment 25 easy words and 25 hard words were presented. In the second half of the experiment the remaining easy and hard words were presented. Each participant was presented with the same words in each half of the experiment, but in a different random order. Halfway through the experiment, participants were given a one-minute rest break. When the experiment resumed, half of the participants heard the same talker present the rest of the stimuli, whereas the other half of the participants heard the other talker present the stimuli. The order of presentation for the talkers was counterbalanced.

When each participant finished the auditory shadowing task, they were asked three questions in the following order:

- (1) Did you notice anything unusual about the experiment?
- (2) Was the first half of the experiment the same as the second half of the experiment?
- (3) Was the voice in the first half of the experiment the same voice that said the words in the second half of the experiment?

These questions were adapted from the naturalistic change blindness experiment by Simons and Levin (1998). Responses to each question were also recorded by the experimenter.

Results

Explicit Measure of Change Deafness

Of the 12 participants who heard the *same* voice in both halves of the experiment, all responded "yes" to question number three, indicating that they had indeed heard the same voice in both halves of the experiment. Of the 12 participants who heard *different* voices in both halves of the experiment, 7 noticed the change in the talker either by stating that the voice was different in response to questions one or two, or by answering "no" to the third question. The remaining 5 participants (42%) did not state that the talker changed when asked questions one and two, and answered "yes" in response to question three, indicating that they failed to detect the change in the talker.

Implicit Measure of Change Deafness

A mixed 2X 3 ANOVA (lexical difficulty as a within factor and talker condition as a between factor) was used to examine the response latencies of the correctly repeated words in the second half of the experiment. Lexical difficulty refers to the easy-hard manipulation among the words. Talker condition was determined by whether a change in talker was presented and if that change was explicitly detected. Listeners who received different talkers in the experiment and failed to explicitly detect the change in the talker are labeled NO in Figure 1. Listeners who received different talkers in the experiment and

explicitly detected the change in the talker are labeled YES in Figure 1. Listeners that received only one talker through out the experiment are labeled SAME in Figure 1.

A main effect of lexical difficulty was found ($F(2, 21) = 24.76, p < .001$) such that easy words (mean = 908 ms) were repeated more quickly than hard words (mean = 934 ms). This result replicates previous studies examining lexical difficulty (e.g., Kirk et al., 1997; Luce & Pisoni, 1998; Sommers, 1996).

The main effect of talker condition was not statistically significant ($F < 1$), but it was in the direction that one might predict based on the results of Goldinger (1998): participants who received the same talker throughout the experiment tended to repeat the words in the second half of the experiment faster than participants who heard different talkers in each half of the experiment. The lack of a significant main effect of talker condition is not unexpected in the present experiment given that *different* easy and hard words were used in each half of the present experiment, whereas Goldinger (1998) used the *same* words in training and test sessions.

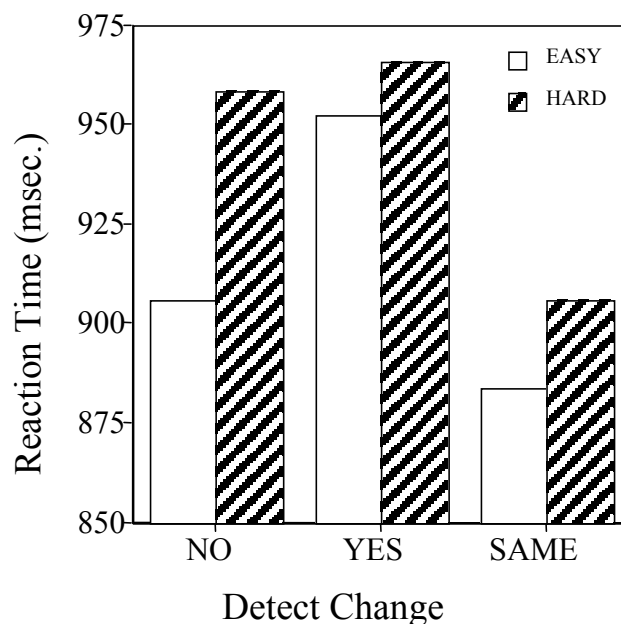


Figure 1. Reaction times from the second half of the experiment to easy and hard words from participants who failed to detect a change in the talker (NO), participants who detected the change in the talker (YES), and participants who received the same talker in both halves of the task (SAME).

Of greatest interest is an interaction between lexical difficulty and talker condition ($F(2, 21) = 3.27, p < .05$). Specifically, the participants that failed to detect the change in the talker had a larger difference between easy and hard words (52 ms) than the participants that detected the change in the talker (13 ms) and the participants that received the same talker throughout the experiment (22 ms). These results suggest that even though participants did not explicitly detect the change in talker, they were implicitly affected by the change in the talker. The results are displayed in Figure 1. No differences in accuracy rates were found (all F 's < 1).

Discussion

The results of the present experiment demonstrate that failures to detect changes occur in the auditory modality as well as the visual modality. Forty-two percent of the participants that heard two different talkers failed to report this change when explicitly questioned about it. Thus, participants may experience “deafness” as well as “blindness” to changes in stimuli. The slightly smaller percentage of participants who failed to detect the change in the present study (42%) compared to other studies of change detection (e.g., 70% in Grimes (1996) and 67% in Experiment 2A of Levin & Simons, 1997) could be due to the differences between auditory speech stimuli and visual stimuli. Speech is a stimulus that is distributed through time (e.g., Marslen-Wilson & Tyler, 1980), whereas visual stimuli are not. Alternatively, the speech used in the present experiment may not have been complex *enough* to be equivalent to the visual stimuli employed in some change detection tasks (e.g., Levin & Simons, 1997). The stimuli used in the present experiment were one-syllable words recorded and presented on high-quality audio equipment with minimal background noise. Perhaps if the words were mixed with noise, the “auditory scene” might be comparable in complexity to the stimuli typically used in visual experiments.

More interestingly, the results of the present experiment demonstrate that even when there was not *explicit* evidence that participants detected the change in talker, there was *implicit* evidence that the change affected processing. Specifically, individuals that were deaf to the change had a greater difference between easy and hard words than participants that detected the change in talkers or participants that had the same talker throughout the experiment. The use of implicit and explicit measures in change detection experiments (see Chun & Nakayama, in press; Hayhoe, in press; Williams & Simons, in press) may provide important insights into cognitive and perceptual processing. For example, work by Nosofsky (1987) suggests that certain stimulus characteristics, or dimensions, of representations in memory can be “stretched” to emphasize a salient aspect. In the present experiment, the difference in reaction times as a function of change detection may be due to different individuals stretching different dimensions of the stimulus to varying degrees. The participants that failed to detect the change in talker may have emphasized the lexical difficulty dimension at the expense of the talker dimension of the spoken words. In contrast, the participants that detected the change in the talker may have equally emphasized the dimensions of talker and lexical difficulty.

Although speculative, this attentional-weighting hypothesis is anecdotally supported by a statement from a participant who failed to detect the change in the talker. When the participant was told during the debriefing of the experiment that there were two different talkers, the participant stated that “I was concentrating so much on *what* he was saying I didn’t pay attention to the voice.” The results of Werner and Thies (in press) also support this hypothesis (see also Shapiro, in press). Werner and Thies found that participants with greater expertise in American football were more likely to detect changes in football images compared to participants with less expertise in American football, suggesting that domain-specific expertise may influence which dimensions of a stimulus are stretched, thereby influencing the detection of changes.

In summary, this experiment demonstrates the existence of *change deafness*. The inability to explicitly detect changes in the auditory domain suggests that change detection is related to attentional demands and is not unique to visual processing. Furthermore, this experiment demonstrates the importance of using *implicit* as well as *explicit* measures of change detection. Although some participants did not explicitly detect the change in the talker, implicit measures of response latency suggest that the change did affect the perceptual and cognitive processing of these participants. Finally, these results support an explanation of change detection based on the distribution of attention-weights and the stretching of stimulus dimensions.

References

- Chun, M.M. & Nakayama, K. (in press). On the functional role of implicit visual memory for the adaptive deployment of attention across scenes. *Visual Cognition*.
- Cohen J.D., MacWhinney B., Flatt M., and Provost J. (1993). PsyScope: A new graphic interactive environment for designing psychology experiments. *Behavioral Research Methods, Instruments, and Computers*, 25, 257-271.
- Goldinger, S.D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological Review*, 105, 251-279.
- Goldinger, S.D., Pisoni, D.B. & Logan, J.S. (1991). On the nature of talker variability effects on recall of spoken word lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 17, 152-162.
- Grimes, J. (1996). On the failure to detect changes in scenes across saccades. In K. Akins (ed.) *Perception: Vancouver Studies in Cognitive Science* (Vol. 2). pp. 89-110. Oxford University Press.
- Hayhoe, M. (in press). Vision using routines: A functional account of vision. *Visual Cognition*.
- Henderson, J.M. & Hollingworth, A. (1999). The role of fixation position in detecting scene changes across saccades. *Psychological Science*, 10, 438-443.
- Kirk, K.I., Pisoni, D.B., Miyamoto, R.C. (1997). Effects of stimulus variability on speech perception in listeners with hearing impairment. *Journal of Speech and Hearing Research*, 40, 1395-1405.
- Kucera, H. & Francis, W.N. (1967). Computational analysis of present-day American English. Providence, RI: Brown University Press.
- Levin, D.T., Momen, N., Drivdahl, S.B. & Simons, D.J. (in press). Change blindness blindness: The metacognitive error of overestimating change-detection ability. *Visual Cognition*.
- Levin, D.T. & Simons, D.J. (1997). Failure to detect changes to attended objects in motion pictures. *Psychonomic Bulletin & Review*, 4, 501-506.
- Luce, P.A. & Pisoni, D.B. (1998). Recognizing spoken words: The neighborhood activation model. *Ear and Hearing*, 19, 1-36.
- Marslen-Wilson, W.D. & Tyler, L. K. (1980). The temporal structure of spoken language understanding. *Cognition*, 8, 1-71.
- Martin, C.S., Mullennix, J.W., Pisoni, D.B. & Sommers, M. (1989). Effects of talker variability on recall of spoken word lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15, 676-684.
- Mullennix, J.W., Pisoni, D.B. & Martin, C.S. (1989). Some effects of talker variability on spoken word recognition. *Journal of the Acoustical Society of America*, 85, 365-378.
- Nosofsky, R.M. (1987). Attention and learning processes in the identification and categorization of integral stimuli. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 13, 87-108.
- Nusbaum, H.C., Pisoni, D.B., & Davis, C.K. (1984). Sizing up the Hoosier mental lexicon: Measuring the familiarity of 20,000 words. *Research on Speech Perception, Progress Report no. 10*. Speech Research Laboratory. Psychology Department, Indiana University, Bloomington, Indiana.
- Nygaard, L.C., Sommers, M. & Pisoni, D.B. (1994). Speech perception as a talker-contingent process. *Psychological Science*, 5, 42-46.
- Palmeri, T.J., Goldinger, S.D., & Pisoni, D.B. (1993). Episodic encoding of voice attributes and recognition memory for spoken words. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 19, 309-328.
- Pisoni, D.B. (1997). Some thoughts on “normalization” in speech perception. In K. Johnson & J.W. Mullennix (Eds.), *Talker Variability in Speech Processing* (pp. 9-32). San Diego: Academic Press.

- Rensink, R.A., O'Regan, J.K. & Clark, J.J. (1997). To see or not to see: The need for attention to perceive changes in scenes. *Psychological Science*, *8*, 368-373.
- Shapiro, K.L. (in press). Change Blindness: Theory or paradigm? *Visual Cognition*.
- Simons, D.J. & Levin, D.T. (1998). Failure to detect changes to people during a real-world interaction. *Psychonomic Bulletin and Review*, *5*, 644-649.
- Sommers, M.S. (1996). The structural organization of the mental lexicon and its contribution to age-related declines in spoken-word recognition. *Psychology and Aging*, *11*, 333-341.
- Torretta, G.M. (1995). The "easy-hard" word multi-talker speech database: An initial report. *Research on Spoken Language Processing, Progress Report No. 20*, Speech Research Laboratory, Psychology Department, Indiana University, Bloomington, Indiana.
- Werner, S. & Thies, B. (in press). Is "change blindness" attenuated by domain-specific expertise? An expert-novices comparison of change detection in football images. *Visual Cognition*.
- Williams, P. & Simons, D.J. (in press). Detecting changes in novel, complex three-dimensional objects. *Visual Cognition*.