THE SOUND OF STROOP: ACOUSTIC EFFECTS IN STROOP INTERFERENCE

Boaz M. Ben-David[1,2,3], Pascal H.H.M. Van Lieshout[1,2,3,4], Alex Nishta[1,3]
[1]Oral Dynamics Lab, Department of Speech-Language Pathology, University of Toronto, St. George Campus, 500 University Avenue, Toronto, Ontario, Canada, M5G 1V7, boaz.ben.david@utoronto.ca
[2]Toronto Rehab, 550 University Avenue, Toronto, Ontario, Canada, M5G 2A2
[3]Centre for Research on Biological Communication Systems, Department of Psychology, University of Toronto Mississauga, 3359 Mississauga Road, N., Mississauga, Ontario, Canada, L5L 1C6
[4]Institute of Biomaterials and Biomedical Engineering, University of Toronto, St. George Campus, 164 College street, Toronto, Ontario, Canada, M5S 3G9

1. INTRODUCTION

Selective attention is a fundamental process in everyday life. In almost every task, one has to attend selectively to certain features in the environment while ignoring or actively suppressing others. For example, talking to a friend in a noisy park, one has to focus on the words of the person sitting next to you on the bench, while ignoring or suppressing conversations taking place around you. If you fail to suppress conversations that surround you, will it affect your speech? Generally, are failures of selective attention manifested in speech?

The most common tool for the assessment of selective attention in the lab has been the Stroop colour-word task (Stroop, 1935). Indeed, the Stroop paradigm has been regarded as the golden standard of selective attention for over 75 years (see Melara & Algom, 2003 for a relevant review). Participants are asked to name the colours of printed words, irrespective of their content. The response latency increase for naming the print colour of an incongruent colour-word (e.g., RED printed in blue) over a colour-neutral stimulus (e.g., OOOO printed in blue) is termed Stroop interference (SI), $SI = CN_{incongruent} - CN_{neutral}$, where CN stands for response latency in colour-naming. If the participant can completely ignore or inhibit the processing of the lexical content of the colored word when asked to name the color in which it was printed, SI should equal zero. On the other hand, if she or he cannot ignore the lexical content of the word, we should observe $SI > 0$.

The SI is taken to reflect an increase in cognitive load and cognitive stress in colour-naming incongruent words (e.g., (Caruso, Chodzko-Zajko, Bidinger, & Sommers, 1994). Failures of selective attention, the semantic conflict presented in an incongruent colour-word, and the process of inhibiting a well-practiced response (reading aloud the printed word) for a less-practiced one (colour-naming) all contribute, in tandem, to the increase in cognitive stress. For example, Renaud and Blondin (1997) found that responding to incongruent stimuli heightened participants' heart rate levels when compared to their heart rate in responding to color-neutral stimuli, indicating higher stress levels.

Stroop effects have been traditionally tested only in terms of response preparation (reaction times and accuracy). However, it stands to reason that these effects go beyond response preparation processes, and have an impact on response production as well, i.e., on the articulation of the oral responses. Specifically, acoustic features of speech have been found to indicate the effect of various emotions and affective conditions, including stress (Scherer, 2003). For example, in the context of our previous example on the conversation in the park, as it takes more effort to distinguish the speech of the person next to you from the other conversations, stress levels are likely to increase as well. This eventually may affect speech, for example, by raising speech intensity.

A few studies (Caruso, & al., 1994; Rothkrantz, Wiggers, Van Wees, & Van Vark, 2004) have investigated features of oral responses in the Stroop paradigm. However, none have actually investigated the acoustic changes in oral responses directly related to the SI. For example, Caruso et al.(1994) compared responses of nine non-pathological participants to incongruent stimuli with responses to congruent ones (RED in red font). However, selective attention fails in both congruent and incongruent trials, facilitating colour responses in the first, while interfering with colour responses in the latter. The results of Caruso et al., therefore, do not indicate the effect of failures of selective attention, but rather an additive effect of interference and facilitation.

In the current study, for the first time, we compared the acoustic characteristics of responses to colour-naming incongruent and neutral stimuli, in an attempt to link the SI in response latency with a parallel Acoustic-SI (ASI), $ASI = AC_{incongruent} - AC_{neutral}$, where AC stands for acoustic characteristics.

2. METHOD

2.1 Participants

Twenty-four young adults participated in the study. All participants were native English speakers, had minimum Snellen fraction (visual acuity test) and pure-tone air-conduction thresholds (from 0.25 to 3.00 kHz) appropriate for their age group. Acoustic data of six participants were not analyzed because of technical problems.

2.2 Procedure and Materials

The study included two blocks of 16 trials, incongruent and neutral. The *incongruent block* included the four colour-words RED, GREEN, BLUE and YELLOW printed in different colours (e.g., RED printed in blue). The *neutral block* included strings of 4 Os printed in green, blue, yellow and red font colours. Participants were tested individually in a single-walled acoustic chamber, seated in front of a computer monitor. In two separate blocks, they were asked to name aloud the font colours of words presented on the monitor speaking into the microphone placed in the table in front of them. In each trial, a single word was presented in the centre of the monitor. After the system registered that a vocal response was initiated, the word disappeared and was replaced by another, one second later.

2.3 Acoustic Analysis.

The digital audio recordings of each participant were split into files that represent single responses. Each response file was compared to the printed stimulus to monitor for the accuracy of the response. Correct responses only were further analyzed using PRAAT software (Boersma, Paul & Weenink, David, 2009). We measured the following acoustic features: fundamental frequency (pitch) and intensity, taken at the midpoint of the vowel; the first, second and third formant frequencies, taken as the mean of values for the entire duration of the vowel; and vowel and word duration. Vowel and word boundaries were manually identified by a trained experimenter while viewing a display showing both a wave form and a broadband spectrogram, with superimposed linear predictive coding based formant tracks (note that in the response "yellow" we analyzed only the first vowel). Acoustic characteristics of single responses were averaged for each experimental block, to allow a within participants comparison design.

3. RESULTS

3.1 Behavioral Data

As commonly found in the literature, accuracy in colour naming was higher for the neutral stimuli than for the incongruent ones, 97.7% and 91%, respectively, $t(23) = 2.72$, $p = 0.01$. A robust SI was found for all participants, with an average SI of 142 ms, $t(23) = 9.9$, $p < 0.001$.

3.2 Acoustic Data

We found significant differences between the acoustic characteristics of responses to incongruent and neutral stimuli (e.g., comparing "blue" responses to RED with "blue" responses to OOOO in blue). Mainly, with incongruent stimuli pitch values decreased, while the intensity and the second and third formant values increased.

4. DISCUSSION

We found significant acoustic differences in colour responses between incongruent and neutral stimuli in the Stroop paradigm. These differences can be specifically linked to the interference caused by the irrelevant task (reading) on colour-naming, unlike previous studies. Finally, our results provide a parallel acoustic effect to the latency effect of Stroop interference. Future research is needed to elucidate the cognitive and affective meaning of these acoustic differences, comparing them to acoustic effects in other attentional tasks.

REFERENCES

Caruso, A. J., Chodzko-Zajko, W. J., Bidinger, D. A., & Sommers, R. K. (1994). Adults who stutter: responses to cognitive stress. *J Psychosom Res, 37,* 746-754.

Boersma, Paul & Weenink, David (2009). Praat: doing phonetics by computer (Version 5.1.10) [Computer program]. Retrieved July 8, 2009, from http://www.praat.org/

Melara, R. D., & Algom, D. (2003). Driven by information: a tectonic theory of Stroop effects. *J Psychosom Res, 110,* 422-471.

Rothkrantz, L., Wiggers, P., Van Wees, J. W., & Van Vark, R. (2004). Voice stress analysis. *Lecture Notes in Artificial Intelligence (Subseries of Lecture Notes in Computer Science), 3206,* 449-456.

Renaud, P., & Blondin, J. P. (1997). The stress of Stroop performance: physiological and emotional responses to color-word interference, task pacing, and pacing speed. *International Journal of Psychophysiol, 27,* 87-97.

Scherer, K. R. (2003). Vocal communication of emotion: A review of research paradigms. *Speech Communication, 40,* 227-256.

Stroop, J. R. (1935). Studies of interference in serial-verbal reaction. *Journal of Experimental Psychology, 18,* 643-662.

ACKNOWLEDGEMENTS

This research was undertaken, in part, thanks to funding from the Canada Research Chairs program and a Research Opportunity Program grant from the Faculty of Arts & Science at the University of Toronto in Mississauga.This study was partially supported by a strategic training grant – Communication and Social Interaction in Healthy Aging, funded by the Canadian Institutes of Health Research. The authors wish to thank the following students for their contribution: Tracey D'Chuna, Huiwen Goy, Wu Yan (Lulu) Li, and Skyler Mooney.