CAPTURING THE REACTION TIME TO DISTINGUISH BETWEEN VOICE AND MUSIC

Alejandro Villena-Rodríguez, Lorenzo J. Tardón, Isabel Barbancho, Ana M. Barbancho, Irene Gómez-Plazas
ATIC Research Group, Universidad de Málaga, Andalucía Tech, E.T.S.I. Telecomunicación,
Dpto. Ingeniería de Comunicaciones, Campus Universitario de Teatinos s/n, 29071, Málaga, Spain
avillenarod@uma.es, lorenzo@ic.uma.es , ibp@ic.uma.es , abp@ic.uma.es , irene_1996f1@uma.es

María-José Varela-Salinas
Universidad de Málaga, Andalucía Tech, Facultad de Filosofía y Letras,
Dpto. Traducción e Interpretación, Campus Universitario de Teatinos s/n, 29071, Málaga, Spain
mjvs@uma.es

ABSTRACT

Reaction times (RTs) are an important source of information in experimental psychology and EEG data analysis. While simple auditory RT has been widely studied, response time when discriminating between two different auditory stimuli have not been determined yet. The purpose of this experiment is to measure the RT for the discrimination between two different auditory stimuli: speech and instrumental music.

1. INTRODUCTION

Reaction time (RT) is defined as the elapsed time between the presentation of a sensory stimulus and the subsequent behavioral response [1]. Simple auditory RT is one of the fastest RTs and is thought to be rarely less than 100 ms [2]. On the other hand, other studies show that the RT, when discriminating between two different stimuli, gets faster as the difference between the stimuli decreases [3].

The experiment considered in this demonstration consist of a go / no-go test where participants are asked to press a button when either a music piece or a speech excerpt is played.

As the stimuli are fairly different in this specific scenario, the RT for the recognition task is expected to be fast but, on the other hand, slower than known simple RTs for auditory stimuli. However, said RT is also expected to increase for speech excerpts presented in a language different from the participant's mother tongue.

2. STIMULI

The stimuli consist of a set of 50 instrumental music and 50 speech excerpts with duration of up to 4 seconds. The music excerpts are randomly cropped from the music tracks shown in Table 1. The presented pieces of speech consists of unconnected sentences recorded in a varied set of languages: Spanish, English, German, French and Danish. In order to keep the variance of volume at the minimum, all the excerpts are compressed and normalized. Each excerpt will be identified by means of pulse width modulation of a trigger/synchronization signal.

Track name	Author
The Magic Flute	Wolfgang Amadeus Mozart
Emperor Waltz	Johann Strauss
Universal Mind	Liquid Tension Experiment
Adios Nonino	Astor Piazzola

Table 1. List of music pieces where the excerpts are taken from.

3. EQUIPMENT

Stimuli will be presented to the subjects through headphones (Sennheiser HD 219) using a system with Presentation software from Neurobehavioral Systems. The response action is recorded by pressing a single USB button connected to the system.

4. METHODOLOGY

In this section, the basis of the methodology employed in the experiments is presented.

4.1 Prior to the experiment

Attendants from the Sound and Music Computing Conference 2019 will be encouraged to participate as experimental subjects. The attendants who will volunteer will be asked to fill in a form with questions about parameters that can have an effect on the response time, such as their gender, age, mother tongue, known languages and conditional factors like fatigue and arousal [4].

Before the experiment begins, a royalty-free track called "Casual Friday" which is unrelated to the experiment will be played, with the objective of setting a comfortable hearing threshold for each participant.

4.2 During the experiment

Once the subject is totally set up, four known test trials will be prompted in order to train the participant. Following these test trials, the music and speech excerpts will be presented to the subject in random order with a pause of 2 seconds between each excerpt.

The experiment, for each subject, will be carried out in 2 separate blocks of trials: in the first block, the subject will be asked to press the button only when a musical piece is played; in the second one, the button should be pressed by the subject only when a speech is played. Both blocks have the same amount of excerpts.

The estimated time for the experiment is 20 minutes plus the explaining and testing time.

5. CONCLUSIONS

The current community lacks of a proper set measurements of the RT when recognizing between speech and music. Also, since the speech excerpts are presented in different languages to a variety of nationalities, the results of this study are not constrained to a certain population with a common mother tongue but expected to be widely valid.

A large set of RTs, like the one described in this demonstration, will not only benefit future EEG studies but will also be helpful to the experimental psychology community or to any researcher interested in understanding how the music and speech is processed in the brain.

Acknowledgments

This work has been funded by the Ministerio de Economía y Competitividad of the Spanish Government under Project No. TIN2016-75866-C3-2-R. This work has been done at Universidad de Málaga, Campus de Excelencia Internacional (CEI) Andalucía TECH.

6. REFERENCES

[1] J. Shelton and G. P. Kumar, "Comparison between auditory and visual simple reaction times," *Neurosci Med*, vol. 1, no. 1, pp. 30–32, 2010.

[2] M. T. Pain and A. Hibbs, "Sprint starts and the minimum auditory reaction time," *Journal of sports sciences*, vol. 25, no. 1, pp. 79–86, 2007.

[3] M. H. Bornstein and N. O. Korda, "Discrimination and matching within and between hues measured by reaction times: Some implications for categorical perception and levels of information processing," *Psychological research*, vol. 46, no. 3, pp. 207–222, 1984.

[4] R. H. Baayen and P. Milin, "Analyzing reaction times," *International Journal of Psychological Research*, vol. 3, no. 2, pp. 12–28, 2010.