Comparing Everyday Listening to Auditory Perception Stimuli by Classifying Two Million YouTube Videos

Andres E. Elizondo Lopez[1,2], & Michael Schutz[3,1,2]

[1]Psychology, Neuroscience and Behavior McMaster University; [2]McMaster Institute for Music and the Mind; [3]School of the Arts McMaster University

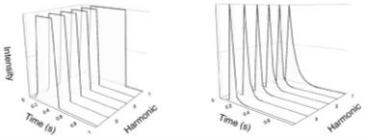
elizonda@mcmaster.ca[1,2]
schutz@mcmaster.ca[3,1,2]

MAPLE Lab maplelab.net

McMaster University

Purpose

- Does sounds affect auditory perception research?
- A large portion of auditory stimuli present in perception experiments are Flat (constant amplitude) (Schutz & Gillard, 2020).
- These flat tones are used in perception experiments to determine the perceptual underpinnings of audition.
- To which degree to which the amplitude envelopes of sounds used in auditory experiments reflect those heard in everyday listening?



[1] Figure 1. A 3D representation of a flat tone, x axis is Time (in seconds), y axis is intensity, z axis is each harmonic component. The flat tone (left) and percussive tone (right).

Social Sciences and Humanities Research Council of Canada | Conseil de recherches en sciences humaines du Canada
Canada
NSERC CRSNG
Canada Foundation for Innovation | Fondation canadienne pour l'innovation

Methods

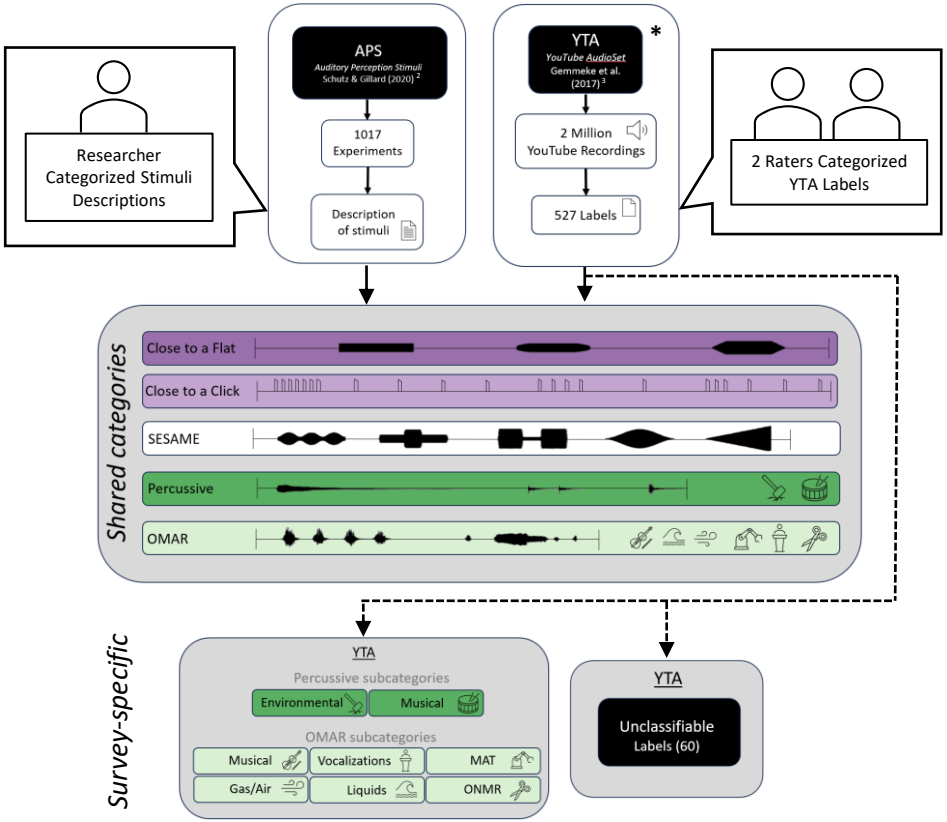
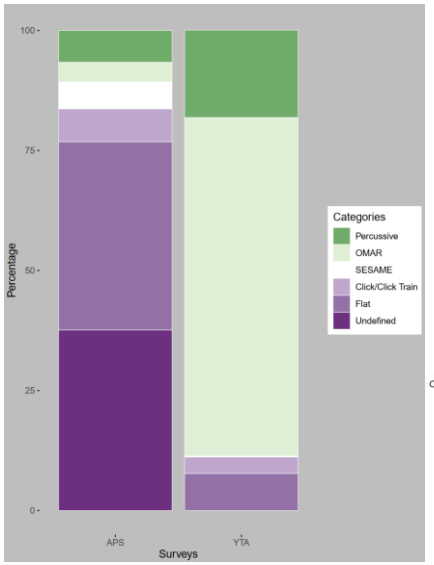


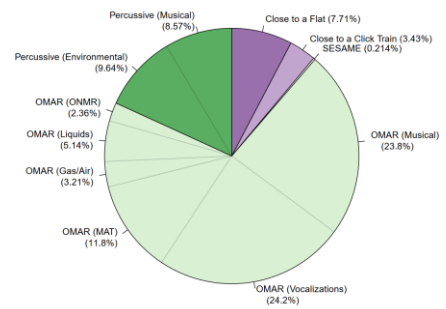
Figure a) represents a flow chart on categorization process. The auditory perception stimuli (APS) represent the stimuli survey investigated by Schutz & Gillard (2020), classifying sounds into different categories from prominent experimental journals.
* The YouTube Audioset (YTA) represents the current study taken from Gemmeke et al . (2017) is a cluster of two million YouTube videos to represent sounds found in everyday encounters, each video are 10 seconds in duration.

Results

b) APS vs YTA Main Category Comparison



c) Subcategories of YTA



d) Weighted by Occurrence Subcategories of YTA

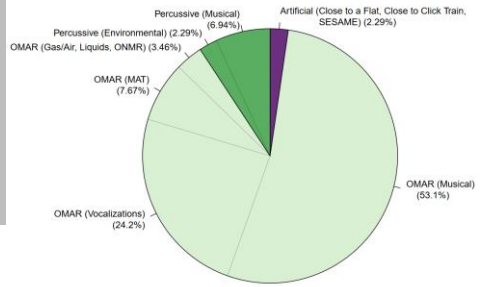


Figure b) represents a bar plot between the auditory perception stimuli (APS) findings from Schutz and Gillard (2020) and the YouTube labels categorized from the YouTube Audioset (YTA). Figure c) represents a pie plot of the subcategories of each label unweighted, figure d) represents how often each label appears in each video

Conclusions

- Majority of the sounds found in the YTA survey fall within the Percussive and OMAR categories.
- Compared to the APS survey, there is a disproportion of categories between sounds found in the laboratory and in everyday encounters.
- This raises questions in the generalization of past findings in auditory perception, as sounds outside of the traditional flat category elicits different results in learning and memory [3], audio-visual integration [4] and duration judgement [5].

Acknowledgements & Selected References

Thank you to Jess LaMantia, J.J. Magee, Max Delle Grazie, Cameron Anderson, Jess Gillard, and the rest of the MAPLE Lab for their help and contributions to the project! Funding provided by NSERC and CFI.

[1] Foley, L., Schlesinger, J., & Schutz, M. (2022). More detectable, less annoying: Temporal variation in amplitude envelope and spectral content improves auditory interface efficacy. The Journal of the Acoustical Society of America, 151(5), 3189–3196.
[2] Gemmeke, J. F., Ellis, D. P. W., Freedman, D., Jansen, A., Lawrence, W., Moore, R. C., Plakal, M., & Ritter, M. (2017). Audio Set: An ontology and human-labeled dataset for audio events. 2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 776–780.
[3] Grassi, M., & Casco, C. (2009). Audiovisual bounce-inducing effect: Attention alone does not explain why the discs are bouncing. Journal of Experimental Psychology: Human Perception and Performance, 35(1), 235–243.
[4] Schutz, M., & Gillard, J. (2020). On the generalization of tones: A detailed exploration of non-speech auditory perception stimuli. Scientific Reports, 10(1), 9520.
[5] Schutz, M., Stefanucci, J. K., Sarah, H. B., & Roth, A. (2017). Name that Percussive Tune: Associative Memory and Amplitude Envelope. Quarterly Journal of Experimental Psychology, 70(7), 1323–1343.
[6] Vallet, G. T., Shore, D. I., & Schutz, M. (2014). Exploring the Role of the Amplitude Envelope in Duration Estimation. Perception, 43(7), 616–630. https://doi.org/10.1068/p7656