

PhD School in Psychology
Department of Cognitive Science
Budapest University of Technology and Economics



Simon Júlia

EXAMINATION OF AUDITORY EVENT SEPARATION

PhD Thesis Booklet

Supervisor: Dr. Winkler István

Budapest, 2022

General background and aim of the thesis

Perception can be defined as the interpretation of accessible information by the cognitive system in a way that influences the present or future behavior of the perceiving organism. Visual stimuli gathered in a single snapshot of time can be rich in details since spacial distance is a major cue in visual object perception. However, auditory perception is mainly based on changes in time (Demany, & Semal, 2008), even loudness perception is dependent on sound duration (Zwislocki, 1969).

Temporal information unfolds on many different time scales, the scale could be 12 orders of magnitude for humans (Mauk & Buonomano, 2004). While the spatial information is computed based on time differences on a microsecond level (Jessell, Kandel, & Schwartz, 2000), a circadian rhythm unfolds during a day.

One of the most popular theoretical approach claims that the processing of incoming stimuli is not continuous, but it is realized in temporal chunks around 30 ms, and events falling in the same window of about 20-40 ms are atemporal, their succession cannot be judged (Pöppel, 1997). Like the resolution of a monitor influences the experience of the image presented, the resolution of the auditory system is assumed to influence higher cognitive processing (Tallal, Merzenich, Miller & Jenkins, 1998; Fink, 2006b; Humes et al, 2013), and differentiates between populations. For example, there is a consensus in the literature that order thresholds tend to be higher in the elderly (Fink, Churan, & Wittmann, 2005; Szymaszek, Szlag, & Sliwowska, 2006; Kolodziejczyk, & Szlag, 2008; Szymaszek, Sereda, Pöppel, & Szlag, 2009; Fogerty, Humes, & Kewley-Port, 2010; Humes, Busey, Craig, & Kewley-Port, 2013). Higher thresholds had been found also in schizophrenia (Capa et al, 2014).

This short temporal window is measured via a Temporal Order Judgment (TOJ) task, in which task participants judge the order of at least two sounds and their threshold is defined based on the inter

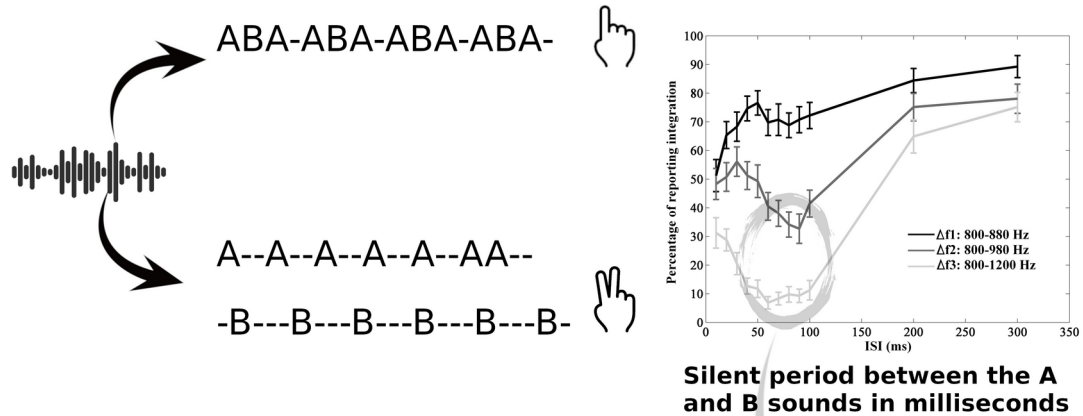
stimuli interval (ISI), where their accuracy reliably drops. First, trained subjects were tested and the threshold was estimated to be around 30 ms, but later studies with large samples lead to the conclusion that the average spatial TOJ threshold is around 50 ms. The spectral version of the task (the two sounds differ in pitch rather than stimulation of different ears) yield to a lower threshold, but subjects can rely on holistic strategies instead of real temporal order judgment (Babkoff, & Fostick, 2013). Based on our results (study 2), even the spatial version of the task is susceptible to distortion, that is why we wanted to explore the average TOJ threshold via electrophysiological measures (study 3 and 4).

Electroencephalography allows the measurement of voltage changes as a result of synchronous activity of neuron populations. The averaged responses that are evoked by a given stimulus type lead to typical positive and negative peak components. The early auditory responses like the P1 (first positive peak), N1 (first negative peak around 100 ms post stimulus), P2 (with latencies around 160-270 ms) are considered obligatory exogenous responses that follow each stimulation. They N1P2 complex is so strongly related to the stimulus, that it can be used in audiometric procedures (Dabbous, El-Shennawy, Hamdy, & Nabieh, 2020).

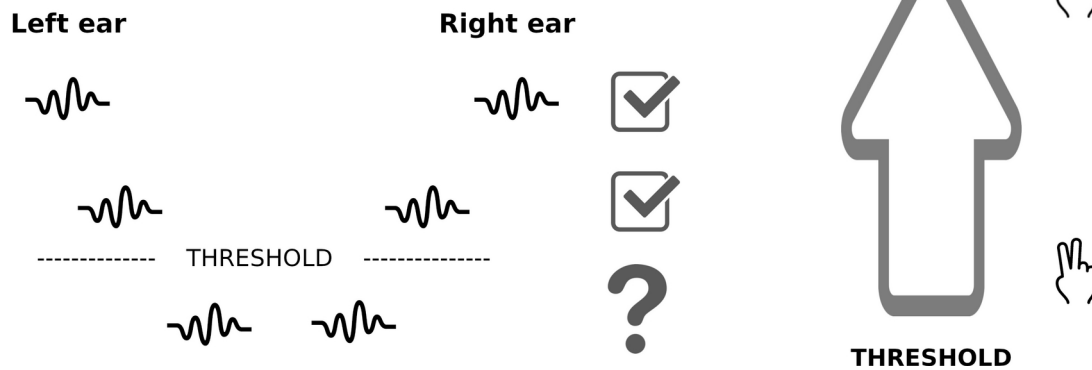
There were attempts to study the electrophysiological correlates of two-tone processing around the temporal order threshold, but the studies are scarce and only one group used the spatial TOJ task (Bernasconi, Grivel, Murray, & Spierer, 2010a, 2010b; Bernasconi, Manuel, Murray, & Spierer, 2011). This group explored the potential correlates of an accurate TOJ judgment, not the electrophysiological correlates of the threshold itself. Therefore, our studies aimed to explore at which ISI can a qualitative change found in the evoked responses, if the time interval between the two sounds are changed in a parametric fashion. Since it can be assumed, that the threshold can be susceptible of contextual effects, we also explored the effect of statistical context (study 3).

New scientific results

Thesis I. The ability to discriminate consecutive auditory events is related to stream segregation at short onset asynchronies



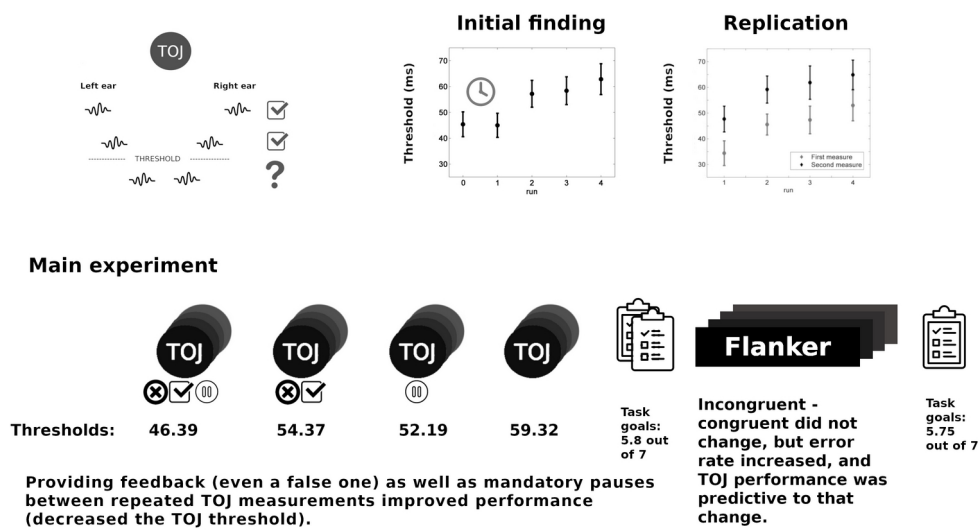
Temporal discrimination measured by the Temporal Order Judgment Task



Inter-individual differences in the chunking mechanism of auditory inputs potentially influence the consecutive grouping of sound events and therefore one's stream segregation ability. In auditory scene analysis (Bregman, 1994), sounds are grouped together based on temporal and/or spectral proximity. In a typical paradigm for studying auditory scene analysis, the ABA0 task, two sounds alternate with the omission of every second B sound. At fast presentation rates, the A sounds and the B sounds form distinguishable auditory streams, while at slow presentation rates only one stream can be heard, that corresponds to a galloping rhythm. Assuming that there is a minimal time interval during which the signals are atemporal, we hypothesized, that at very high presentation rates, the general tendency to hear separate streams inverts and the smaller the ISI between the consecutive sounds, the more likely it will be that listeners report hearing the integrated percept. Also, we

expected to find correlations between one's ability to discriminate consecutive sounds measured by the spatial TOJ task and their ability to hear segregated streams at very fast presentation rates. These assumptions were supported by our results, as at the fastest rates, the ratio of integrated/segregated percept reports did increase. In addition, just around the ISI of the inversion, the percept ratios and the TOJ thresholds did show a significant correlation.

Thesis II. The Temporal Order Judgment Task is susceptible to fatigue

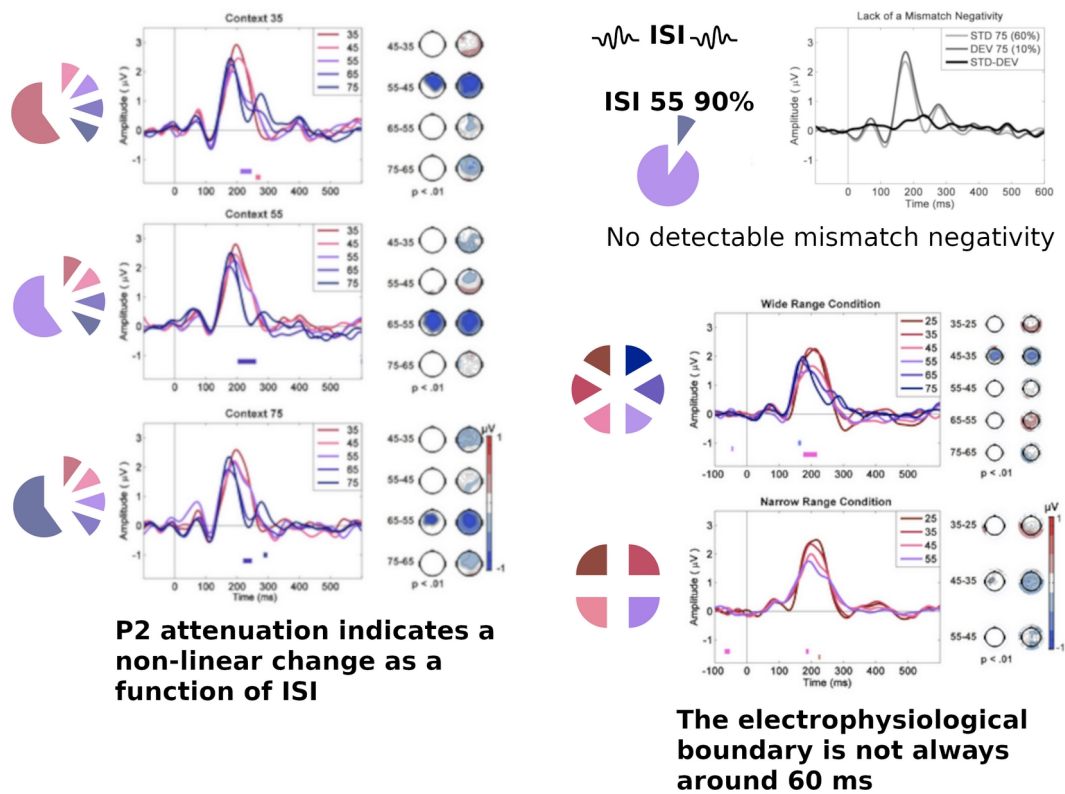


The TOJ task not only measures one's discrimination ability, but also the ability to correctly report the sound sequence. It can be assumed that the task requires higher order processes, such as attention, and hence it is susceptible to cognitive fatigue. Our previous findings were surprising in the respect that repeating the threshold measurement only four times in a row already led to performance deterioration. We wanted to replicate this result and also show that the effect can be eliminated by mandatory pauses between measurements and also by a motivating feedback (comparing the performance to a reference value). The performance deterioration was replicated multiple times and even the subjective fatigue increased during the experiments. However, we were not able to find correlation between the performance deterioration and subjective fatigue. One possible reason is that subjective fatigue is a result of cost-benefit evaluation that is not strictly related to the actual depletion of cognitive resources. Our results showed that subjective fatigue

could be predicted by the change in positive affectivity, by the experienced fatigue at the beginning of the experiment, or by one's sensitivity to criticism.

It was interesting that the TOJ threshold was predictive of the attentional performance but not vice versa. This means that attentional performance is either related to the temporal discrimination ability of the participant or that the TOJ threshold is highly sensitive even to the small lapses in attention, more so than the applied attention task (flanker).

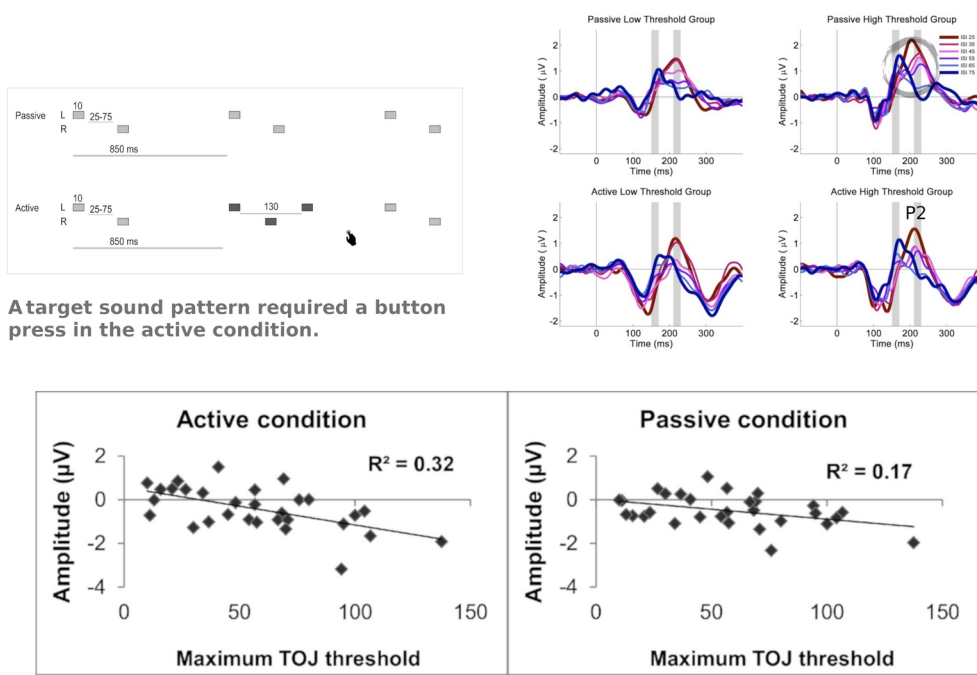
Thesis III. Event-Related Potentials to tone pairs show a boundary-like effect at ca. 60 ms within-pair ISI



If there is a qualitative change between processing two sounds as one auditory event versus two successive events, then the parametric manipulation of the ISI should reveal a non-linear change in the electrophysiological response to the tone pairs. While recording EEG, we presented tone pairs differing in their within-pair ISI (from 25 ms to 75 ms in 10 ms steps) in a pseudorandom order during a passive listening condition when participants were attending a silent movie with subtitles.

The probability of tone pairs having short or long ISIs was also manipulated in order to examine the flexibility of a possible event boundary. Comparing differences of ERPs of tone pairs differing in neighboring ISI values a qualitative change was revealed between the ERPs of tone pairs having 55 and 65 ms of ISIs. This electrophysiological marker was a negative component in the time interval of the P2 component and it appeared earlier (between 45 and 55 ms) when the short ISI pairs were presented in higher proportion. However, we were not able to find correlation between the participants' behavioral TOJ threshold and the ERP difference wave where the qualitative change occurred.

Thesis IV. The pronounced amplitude change at about 60 ms is related to behavioral performance in the Temporal Order Judgment task



More pronounced P2 attenuation between ERPs having ISI 55 ms and 65 ms predicts a higher TOJ threshold.

Based on the conclusions of Thesis II, it can be assumed that the behavioral TOJ threshold is highly sensitive to the actual state of the participant, and that makes it hard to find correlation between the electrophysiological markers obtained in a passive condition and the behavioral threshold, which requires strong attentional focus. Therefore, an active condition was designed to measure the electrophysiological correlates of boundary crossing (from fused to double events as a function of

ISI). The inferred ISI at which the transition from perceiving one to two events occurred was not calculated as it did not seem sufficiently reliable, and instead the amplitude of the negative difference between the ERPs of tone pairs having an ISI = 55 ms and ISI = 65 ms was measured. Other adjacent ERP differences were not as relevant as at short ISIs where almost everyone hears fused tones, while at the longest ISIs most people hear successive tones. This approach turned out to be valid, as a significant medium-size correlation was found between the behavioral TOJ threshold and the amplitude decrease in the P2 region between the ERPs elicited by tone pairs having an ISI = 55 ms and ISI = 65 ms. This correlation was present even in the passive condition, but with a smaller effect size.

Conclusions

Individual differences in temporal discrimination measured via the spatial Temporal Order Judgment task can be in association with auditory stream segregation, therefore one's ability to extract the relevant auditory object when concurrent streams are present. Based on our first study, alternating sounds having an onset difference around the TOJ threshold tend to be perceived as one stream by individuals having a high TOJ threshold. However, the measured relationship was not strong, but that could be explained by the low test-retest reliability of the spatial TOJ threshold, since it is susceptible to many factors. For example, our incidental finding showed that it gets worse with repeated consecutive measurements.

In the second study, we presented more than one replication of the above mentioned incidental finding. However, the performance deterioration could have been diminished via pauses and via feedback. These results point to the interpretation that this task is highly susceptible even to short term cognitive fatigue. Normally, even hours are required to reliably measure performance degradation (Takács et al, 2019). The quick performance decline and the subjective fatigue in this task can be explained by a high cost/benefit ratio as the task is monotonic, not rewarding. Another possible explanation is that the adaptive algorithm that we use is susceptible even for the short

lapses of attention that is not detectable with attention tasks comparing block level changes (averages of longer periods). Subjective fatigue also increased during the experiment, but correlation could not be found. Subjective fatigue was related to decrease in positive affectivity, perceived difficulty or effort, sensitivity to criticism. These results suggest that, under normal circumstances, the feeling of fatigue is more related to the assessment of the given situation than to the actual depletion of cognitive resources.

Since the behavior task is susceptible to many factors, we wanted to find an objective measure of the threshold. Our ERP studies revealed a qualitative change in the ERP as a function of ISI. A pronounced P2 attenuation was found around 60 ms. This threshold proved to be flexible, it was influenced by the statistical context and experimental setup. In our fourth study, we could show a correlation between the P2 attenuation and the behavioral threshold, that tended to be higher in the attended condition, when the evoked responses were expected to be less variable.

It can be concluded that there is indeed a physiologically and behaviorally relevant temporal constrain in auditory perception around 40-60 ms, that is hard to measure only via behavioral tasks, but the attenuation of the P2 ERP component can be informative and hopefully be a promising tool in understanding individual differences in auditory processing. Also, the spatial TOJ task might be a quick tool in studying the various factors that interact influencing subjective as well as cognitive fatigue.

List of publications related to the theses

- I. Simon, J., & Winkler, I. (2018). The role of temporal integration in auditory stream segregation. *Journal of Experimental Psychology: Human Perception and Performance*, 44(11), 1683.
- II. Simon, J., Takács, E., Orosz, G., Berki, B., & Winkler, I. (2020). Short-term cognitive fatigue effect on auditory temporal order judgments. *Experimental brain research*, 238(2), 305-319.
- III. Simon, J., Balla, V., & Winkler, I. (2019). Temporal boundary of auditory event formation: An electrophysiological marker. *International Journal of Psychophysiology*, 140, 53-61.
- IV. Júlia, S., & Balla, V. R. (2020). Electrophysiological correlates of the spatial temporal order judgment task. *Biological Psychology*, 156, 107947.

References

- Babkoff, H., & Fostick, L. (2013). The role of tone duration in dichotic temporal order judgment. *Attention, Perception, & Psychophysics*, 75(4), 654-660.
- Bernasconi, F., Manuel, A., Murray, M., & Spierer, L. (2011). Pre-stimulus beta oscillations within left posterior sylvian regions impact auditory temporal order judgment accuracy. *International Journal of Psychophysiology*, 79(2), 244-248. doi:10.1016/j.ijpsycho.2010.10.017
- Bernasconi, F., Grivel, J., Murray, M., & Spierer, L. (2010a). Plastic brain mechanisms for attaining auditory temporal order judgment proficiency. *NeuroImage*, 50(3), 1271-1279. doi:10.1016/j.neuroimage.2010.01.016
- Bernasconi, F., Grivel, J., Murray, M., & Spierer, L. (2010b). Interhemispheric coupling between the posterior sylvian regions impacts successful auditory temporal order judgment. *Neuropsychologia*, 48(9), 2579-2585.
- Bregman, A. S. (1994). *Auditory scene analysis: The perceptual organization of sound*. MIT press.
- Capa, R., Duval, C., Blaison, D., & Giersch, A. (2014). Patients with schizophrenia selectively impaired in temporal order judgments. *Schizophrenia Research*, 156(1), 51-55. doi:10.1016/j.schres.2014.04.001
- Dabbous, A. O., El-Shennawy, A. M., Hamdy, M. M., & Nabieh, S. F. (2020). COMPARISON OF N1P2 CORTICAL AUDITORY EVOKED POTENTIAL AND NARROW-BAND CHIRP AUDITORY STEADY STATE POTENTIAL IN HEARING THRESHOLD DETECTION IN ADULTS. *Journal of Hearing Science*, 10(4).
- Demany, L., & Semal, C. (2008). The role of memory in auditory perception. In *Auditory perception of sound sources* (pp. 77-113). Springer, Boston, MA.

- Fink, M., Ulbrich, P., Churan, J., & Wittmann, M. (2006b). Stimulus-dependent processing of temporal order. *Behavioural processes*, 71(2-3), 344-352.
- Fink, M., Churan, J., Wittmann, M. (2005). Assessment of auditory temporal-order thresholds – A comparison of different measurement procedures and the influences of age and gender. *Restor Neurol Neurosci*, 23, 281–296
- Fogerty, D., Humes, L. E., & Kewley-Port, D. (2010). Auditory temporal-order processing of vowel sequences by young and elderly listeners. *The Journal of the Acoustical Society of America*, 127(4), 2509-2520.
- Humes, L. E., Busey, T. A., Craig, J., & Kewley-Port, D. (2013). Are age-related changes in cognitive function driven by age-related changes in sensory processing?. *Attention, Perception, & Psychophysics*, 75(3), 508-524.
- Jessell, T. M., Kandel, E. R., & Schwartz, J. H. (2000). *Principles of neural science* (No. 577.25 KAN).
- Kolodziejczyk, I., & Szlag, E. (2008). Auditory perception of temporal order in centenarians in comparison with young and elderly subjects. *Acta Neurobiol Exp*, 68(3), 373-81.
- Mauk, M. D., & Buonomano, D. V. (2004). The neural basis of temporal processing. *Annual Review of Neuroscience*, 27, 307–340.
- Pöppel, E. (1997). A hierarchical model of temporal perception. *Trends in cognitive sciences*, 1(2), 56-61.
- Szymaszek, A., Szlag, E., Sliwowska, M. (2006). Auditory perception of temporal order in humans: The effect of age, gender, listener practice and stimulus presentation mode. *Neurosci Lett*, 403, 190–194

- Szymaszek, A., Sereda, M., Pöppel, E., Szélag, E. (2009). Individual differences in the perception of temporal order: The effect of age and cognition. *Cogn Neuropsychol*, 26(2), 135–147.
- Takács E, Barkaszi I, Altbäcker A, Czigler I, Balázs L (2019) Cognitive resilience after prolonged task performance: an ERP investigation. *Exp Brain Res* 237(2):377–388
- Tallal, P., Merzenich, M., Miller, S., Jenkins, W., 1998. Language learning impairments: integrating basic science, technology, and remediation. *Exp. Brain Res.* 123, 210–219.
- Zwislocki, J. J. (1969). Temporal summation of loudness: An analysis. *The Journal of the Acoustical Society of America*, 46(2B), 431-441.