WHAT CAUSES THE PRIMACY EFFECT IN TEMPORAL LOUDNESS WEIGHTS?

Daniel Oberfeld[1] **Alexander Fischenich[1]** **Emmanuel Ponsot[2,3]**
Jesko L. Verhey[4] **Jan Hots[4]**
[1] Experimental Psychology, Johannes Gutenberg-Universität Mainz, Germany
[2] Hearing Technology @ WAVES, Department of Information Technology, Ghent University, Belgium
[3] Laboratoire des systèmes perceptifs, Département d'études cognitives, École normale supérieure PSL University, France
4 Experimental Audiology, Otto-von-Guericke-Universität Magdeburg, Germany
oberfeld@uni-mainz.de

ABSTRACT

Studies on loudness judgments for time-varying sounds consistently show a primacy effect: The beginning of a sound is more important for the judgment of its overall loudness than later temporal parts. The mechanisms that were proposed as an explanation of this pattern of temporal loudness weights encompass a wide range of concepts. Ordering the proposed explanations from early auditory mechanisms to supra-modal cognitive processes, the primacy effect has been linked to a) the onset peak of auditory nerve fibers, b) effects of non-simultaneous masking on auditory intensity resolution, c) capture of auditory attention to the sound onset, d) temporal integration of evidence at a decisional stage (sequential sampling), and e) serial position effects in a memory system. This talk discusses each of the alternative concepts in the light of the available data, and identifies gaps in the literature that need to be filled in order to decide between the alternatives. Also, the predictions of current loudness models concerning temporal weights are evaluated.

1. TEMPORAL WEIGHTS IN LOUDNESS JUDGMENTS

Because loudness is one of the fundamental aspects of auditory sensation and is important when it comes to the perception of our environment through the auditory channel, extensive research on steady-state sounds has been done in the past and loudness models are available that account for a large proportion of the psychoacoustic data [1-3]. However, a substantial number of studies showed that not all temporal parts of a sound are weighted equally when listeners judge loudness, and that the perception of time-varying sounds can thus not be fully accounted for by static loudness models [4, 5]. The reported perceptual weights assigned to the different temporal parts of a sound show consistently that the beginning of a time-varying sound is of higher importance for the perception of loudness of the sound as a whole ("global loudness") than later temporal parts [e.g., 6-9, 10], which has been referred to as a *primacy effect* (for an overview, see [11]). The primacy effect can be described by an exponential decay function with a time-constant of about 200-300 ms [11, 12]. Across experiments, the temporal weight at the beginning of the sound is 4 to 5 times higher than the asymptotic weight, and the weight assigned to a temporal portion of a sound is the integral of this function over the segment duration [12].

2. PROPOSED EXPLANATIONS

The mechanisms that were proposed as an explanation of this pattern of temporal loudness weights encompass a wide range of concepts. If we order the proposed explanations from early auditory mechanisms to supra-modal cognitive processes, the first potential origin of the primacy effect, proposed in [10], is **a) the onset peak in the firing rate of auditory nerve (AN) fibers** [13]. If loudness is assumed to be related to the spike count elicited by the sound [14], and the onset causes more neural activity than later temporal parts, this could explain a higher loudness weight on the sound onset. As the inner hair cells that innervate the AN fibers are frequency specific, the recovery of the firing rate is also frequency specific [e.g., 15]. This is compatible with results demonstrating that temporal loudness weights are applied in a frequency-specific manner [16], and that when the spectrum changes abruptly within a contiguous sound, a second primacy effect is observed on the second sound part [8]. However, because neurons with high spontaneous rates (SR) show a fast recovery, so that the onset peak occurs after silent inter-stimulus intervals of only a few milliseconds [15], the observation that the primacy effect in loudness judgments shows full recovery only after silent gaps of about 350 ms or more [17] seems, at first sight, to be at odds with this explanation. Yet, low-SR neurons exhibit a considerably slower recovery of the onset peak [18, 19], thus one could assume that the primacy effect is primarily driven by these neurons. Another result that is not easily accounted for by the onset peak in AN fibers is that varying the mean sound level from just above detection threshold to higher levels, or presenting to-be-judged sound in a continuous background noise has almost no effect on the temporal weights [20], while the AN responses are strongly influenced by sound level and simultaneous masking. On a more general level, the neuronal auditory pathway is quite complex and involves different types of neurons as well as efferent and afferent loops. Additional research based on predictions from computational auditory models is thus needed to evaluate

10.48465/fa.2020.0835

more exactly the extent to which processes in the auditory periphery might to the primacy effect in loudness weights. Second, **b) effects of non-simultaneous masking on auditory intensity resolution** were proposed as a potential origin of the primacy effect in loudness judgments [21]. The intensity resolution for a sound can be severely impaired by more intense sounds presented shortly before the to-be-judged sound [e.g., 22, 23]. For a level fluctuating sound, the probability that a given temporal segment is forward-masked by a temporal segment higher in level increases with the number of segments that precede the actual portion. Therefore, on average, the intensity resolution for later temporal parts of a sound might be reduced relative to earlier parts. If now the listeners follow an "ideal observer" strategy of placing higher weights on temporal portions of a sound for which the intensity resolution is high [24, 21], this would result in higher weights assigned to the beginning of the sound. As the masking effect on intensity resolution is frequency-specific [25], this explanation can account for corresponding results [8, 16]. Also, the masking effect is reduced when the inter-stimulus interval between masker and target sound increases [25], compatible with the recovery of the primacy effect [17]. In addition, the observation that stimulus components with, on average, higher loudness than the remaining components receive higher weights ("loudness dominance") [26-28] is compatible with this explanation, because sound parts higher in level can have a masking effect on weaker elements, but not vice versa [23]. However, a serious limitation of this account is that sounds presented *after* rather than *before* a target sound can also reduce the intensity resolution [29, 30]. Using the same rationale as for forward masking, backward masking effects would result in the *opposite* temporal weighting pattern compared to forward masking effects, with pronounced *recency effects* (i.e., higher weights at the end of a sound) and lower weights at the onset of the sound, because the probability that a segment is followed by a segment higher in level decreases with the serial position of the segment within the sound. While one study reported the backward masking effects to level off more quickly than forward masking when the masker-target interval became longer [29], additional experiments are required. In particular, it remains to be determined whether the asymmetry between forward and backward masking effects is in fact sufficient to explain the consistent observation of a consistent *primacy* effect, rather than a recency effect, in temporal loudness weights.

Third, as proposed in [10], the primacy effect might originate from **c) an attention orientation response [31-33] to the sound onset**, in the sense that the sound onset captures the attention [34]. Higher attention assigned to the beginning rather than to later parts of the sounds corresponds to higher perceptual weights at sound onset [35]. However, the data in [6] show that sounds starting with a gradual increase in level across the first few hundred milliseconds (i.e. "fade-in") lead to a "delayed primacy effect", in the sense of a high weight on the first unattenuated temporal segment. This is incompatible with the attentional explanation, because the gradual increase in level should have reduced the perceived abruptness of sound onset and, as a consequence, the capture of attention [36, 37]. Related, presenting the to-be-judged sounds only a few dB above the detection threshold in a continuous background noise, which should also reduce the perceived abruptness of the target sound onset, did not result in a substantial reduction in the primacy effect [16]. Thus, attentional capture cannot fully account for primacy effects in loudness judgments, although some aspects like the effect of a sudden change in sound spectrum [8] or the effect of silent gaps within the sound [17] are compatible with this explanation.

Fourth, the primacy effect might represent a decisional rather than a sensory effect. As proposed in [20], it might be the result of **d) an "evidence integration" or "sequential sampling" process** in which evidence for one of the two possible alternatives (e.g., is the sound louder or softer than previous sound in the experimental block) is collected sequentially during each trial and the decision is made as soon as sufficient information has been accumulated, ignoring further evidence during the trial (e.g., [38, 39]). The reoccurrence of a primacy effect after a silent gap [17] or after a sudden change in the frequency spectrum [8] can be accounted for by assuming that the sound parts before and after the gap, or before and after the spectral change, are perceived as two separate sounds or auditory objects rather than as one unitary sound. In this case, listeners might make a separate judgment (i.e., start separate evidence integration processes) of the loudness for each of the two sound parts. Following the same line of reasoning, if one assumes that a separate evidence integration process is executed for each frequency band or auditory stream, then frequency-specific weights are predicted [16]. An interesting implication of such a decisional account for the primacy effect is that similar temporal weighting patterns should occur in different sensory modalities, because the decision process is supramodal. In fact, primacy effects have been reported for visual motion perception [39] or brightness perception [40], using time-varying stimuli as in the loudness judgments tasks. In a recent study [41], we tested the prediction based on this account of the primacy effect that the temporal weights applied in 1) a loudness judgment task for sounds varying in level across time, and 2) a brightness judgment task for visual stimuli varying in brightness across time should be similar (provided that the temporal fluctuations and the difficulty of the tasks are comparable), because both of them are due to the same evidence integration process at a supra-modal decision stage. The results showed a primacy effect in the loudness judgments, but not in the brightness judgments, and thus did not support the hypothesis. However, additional research is needed in order to identify the extent to which the deviation between the loudness and brightness weights can be attributed to differences in sensory processing, such as afterimage effects in the visual task.

Fifth, in [9] it was proposed that the primacy effect in loudness judgments might represent **e) a serial position effect in a memory system**. Primacy and recency effects are ubiquitous in experiments on short-term memory [e.g., 42, 43]. One could assume that the sequence of sound segments of a time-varying sound is stored and processed in a memory system exhibiting similar characteristics to verbal short-term memory. If this was the case, the memory representation of the first few segments should be more precise or easier accessible than for the remaining segments, causing a primacy effect when listeners combine the sequence of segment loudness representations into a global loudness judgment for the sound as a whole. However, in memory tasks, the primacy effects typically encompass approximately the first 3-5 items in the list, and are relatively independent of the item duration [42, 44]. Put differently, the recognition accuracy is a function of the relative position of the item within the list (e.g., first item or middle item), roughly independent of the time delay between list onset and item onset. Exactly the opposite pattern is observed in temporal loudness weights, where the weight assigned to a given temporal segment is determined by the time delay between sound onset and segment onset, roughly independent of the sound duration [11]. In addition, the primacy effect in memory does not decay within less than one second as it does for the temporal loudness weights [12], and serial position curves in memory often additionally show a clear recency effect, or even only a recency but no primacy effect [e.g., 42, 45]. Thus, the primacy effect in loudness behaves very differently than serial position effects in memory.

Finally, **loudness models for dynamic, time-varying sounds** have been proposed [46-49], including temporal integration stages. However, these models do not predict the observed primacy effects in loudness judgments [50, 20, 51]. Therefore, first attempts were made to extend existing models to account for the observed temporal weighting [51].

3. CONCLUSION

Except for the explanations based on capture of attention (**c**) and serial position effects in memory (**e**), all of the other potential explanations of the primacy effect in loudness judgments discussed above (**a**, **b**, and **d**) are compatible with some important aspects of the data, while none of them appears to be able to predict the entire pattern of results on temporal loudness weights. Some aspects of the results may already be understood based on auditory nerve responses, but in our view, all aspects can only be accounted for by assuming that higher-level processes also contribute to loudness weights. Additional research is needed to clarify which mechanism(s) underlie the observed temporal loudness weights, as well as their respective contribution.

4. REFERENCES

[1] Scharf, B., "Loudness", *Handbook of Perception, Volume IV. Hearing*, E.C. Carterette and M.P. Friedman, Editors. 1978, Academic Press: New York. p. 187–242.

[2] Glasberg, B.R., and B.C. Moore, "Prediction of absolute thresholds and equal-loudness contours using a modified loudness model". *Journal of the Acoustical Society of America*, 2006. **120**(2): p. 585-8.

[3] Jesteadt, W., and L. Leibold, "Loudness in the Laboratory, Part I: Steady-State Sounds", *Loudness*, M. Florentine, A.N. Popper, and R.R. Fay, Editors. 2011, Springer. p. 109-144.

[4] Rennies, J., J.L. Verhey, J.E. Appell, and B. Kollmeier, "Loudness of complex time-varying sounds ? A challenge for current loudness models". *Proceedings of Meetings on Acoustics*, 2013. **19**.

[5] Rennies, J., M. Wächtler, J. Hots, and J.L. Verhey, "Spectro-temporal characteristics affecting the loudness of technical sounds: data and model predictions". *Acta Acustica United with Acustica*, 2015. **101**(6): p. 1145-1156.

[6] Namba, S., S. Kuwano, and T. Kato, "Loudness of sound with intensity increment". *Japanese Psychological Research*, 1976. **18**(2): p. 63-72.

[7] Ellermeier, W., and S. Schrödl, "Temporal weights in loudness summation", *Fechner Day 2000. Proceedings of the 16th Annual Meeting of the International Society for Psychophysics*, C. Bonnet, Editor. 2000, Université Louis Pasteur: Strasbourg. p. 169-173.

[8] Pedersen, B., and W. Ellermeier, "Temporal weights in the level discrimination of time-varying sounds". *Journal of the Acoustical Society of America*, 2008. **123**(2): p. 963-972.

[9] Dittrich, K., and D. Oberfeld, "A comparison of the temporal weighting of annoyance and loudness". *Journal of the Acoustical Society of America*, 2009. **126**(6): p. 3168-3178.

[10] Oberfeld, D., and T. Plank, "The temporal weighting of loudness: Effects of the level profile". *Attention, Perception, & Psychophysics*, 2011. **73**(1): p. 189-208.

[11] Oberfeld, D., J. Hots, and J.L. Verhey, "Temporal weights in the perception of sound intensity: Effects of sound duration and number of temporal segments". *Journal of the Acoustical Society of America*, 2018. **143**(2): p. 943-953.

[12] Oberfeld, D., L. Jung, J.L. Verhey, and J. Hots, "Evaluation of a model of temporal weights in loudness judgments". *Journal of the Acoustical Society of America*, 2018. **144**(2): p. EL119-EL124.

[13] Kiang, N.Y.S., T. Watanabe, E.C. Thomas, and L.F. Clark, *Discharge patterns of single fibers in the cat's auditory nerve*. 1965, Cambridge, Mass.: M.I.T. Press.

[14] Relkin, E.M., and J.R. Doucet, "Is loudness simply proportional to the auditory nerve spike count?". *Journal of the Acoustical Society of America*, 1997. **101**(5 Pt 1): p. 2735-40.

[15] Harris, D.M., and P. Dallos, "Forward masking of auditory-nerve fiber responses". *Journal of Neurophysiology*, 1979. **42**(4): p. 1083-1107.

[16] Fischenich, A., J. Hots, J.L. Verhey, and D. Oberfeld, "Temporal loudness weights are frequency specific". Manuscript under review.

[17] Fischenich, A., J. Hots, J.L. Verhey, and D. Oberfeld, "The effect of silent gaps on temporal

weights in loudness judgments". *Hearing Research*, 2020. **395**.
[18] Rhode, W.S., and P.H. Smith, "Characteristics of tone-pip response patterns in relationship to spontaneous rate in cat auditory nerve fibers". *Hearing Research*, 1985. **18**(2): p. 159-168.
[19] Relkin, E.M., and J.R. Doucet, "Recovery from prior stimulation. I: Relationship to spontaneous firing rates of primary auditory neurons". *Hearing Research*, 1991. **55**(2): p. 215-222.
[20] Fischenich, A., J. Hots, J. Verhey, and D. Oberfeld, "Temporal weights in loudness: Investigation of the effects of background noise and sound level". *PloS one*, 2019. **14**(11).
[21] Oberfeld, D., M. Kuta, and W. Jesteadt, "Factors limiting performance in a multitone intensity-discrimination task: Disentangling non-optimal decision weights and increased internal noise". *PLOS One*, 2013. **8**(11): p. e79830.
[22] Zeng, F.G., C.W. Turner, and E.M. Relkin, "Recovery from prior stimulation II: Effects upon intensity discrimination". *Hearing Research*, 1991. **55**(2): p. 223-230.
[23] Oberfeld, D., "The mid-difference hump in forward-masked intensity discrimination". *Journal of the Acoustical Society of America*, 2008. **123**(3): p. 1571-1581.
[24] Green, D.M., "Detection of multiple component signals in noise". *Journal of the Acoustical Society of America*, 1958. **30**(10): p. 904-911.
[25] Zeng, F.G., and C.W. Turner, "Intensity discrimination in forward masking". *Journal of the Acoustical Society of America*, 1992. **92**(2): p. 782-787.
[26] Lutfi, R.A., and W. Jesteadt, "Molecular analysis of the effect of relative tone level on multitone pattern discrimination". *Journal of the Acoustical Society of America*, 2006. **120**(6): p. 3853-60.
[27] Oberfeld, D., "Does a rhythmic context have an effect on perceptual weights in auditory intensity processing?". *Canadian Journal of Experimental Psychology-Revue Canadienne De Psychologie Experimentale*, 2008. **62**(1): p. 24-32.
[28] Ponsot, E., P. Susini, G. Saint Pierre, and S. Meunier, "Temporal loudness weights for sounds with increasing and decreasing intensity profiles". *Journal of the Acoustical Society of America*, 2013. **134**(4): p. EL321-EL326.
[29] Plack, C.J., R.P. Carlyon, and N.F. Viemeister, "Intensity discrimination under forward and backward masking: role of referential coding". *Journal of the Acoustical Society of America*, 1995. **97**(2): p. 1141-9.
[30] Oberfeld, D., and P. Stahn, "Sequential grouping modulates the effect of non-simultaneous masking on auditory intensity resolution". *PLoS ONE*, 2012. **7**(10): p. e48054.
[31] Pavlov, I.P., *Conditioned reflexes: An investigation of the physiological activity of the cerebral cortex. Translated and edited by G. V. Anrep*. 1927, London: Oxford University Press.
[32] Sechenov, I.M., *Reflexes of the brain (original publication 1863)*. 1965, Cambridge, Mass.: M.I.T. Press.
[33] Graham, F.K., and S.A. Hackley, "Passive and active attention to input", *Handbook of cognitive psychophysiology: Central and autonomic nervous system approaches*, J.R. Jennings and M.G.H. Coles, Editors. 1991, Wiley: Chichester, West Sussex, England; New York, NY, USA. p. 251-356.
[34] Koelewijn, T., A. Bronkhorst, and J. Theeuwes, "Auditory and visual capture during focused visual attention". *Journal of Experimental Psychology: Human Perception and Performance*, 2009. **35**(5): p. 1303-1315.
[35] Berg, B.G., "Observer efficiency and weights in a multiple observation task". *Journal of the Acoustical Society of America*, 1990. **88**(1): p. 149-58.
[36] Jonides, J., and S. Yantis, "Uniqueness of abrupt visual onset in capturing attention". *Perception and Psychophysics*, 1988. **43**(4): p. 346-354.
[37] Ruz, M., and J. Lupiáñez, "A review of attentional capture: On its automaticity and sensitivity to endogenous control. ". *Psicológica*, 2002. **23**(2): p. 283-309.
[38] Vickers, D., "Evidence for an accumulator model of psychophysical discrimination". *Ergonomics*, 1970. **13**(1): p. 37-58.
[39] Tsetsos, K., J. Gao, J.L. McClelland, and M. Usher, "Using time-varying evidence to test models of decision dynamics: Bounded diffusion vs. the leaky competing accumulator model". *Frontiers in Neuroscience*, 2012. **6**.
[40] Bronfman, Z.Z., N. Brezis, and M. Usher, "Non-monotonic temporal-weighting indicates a dynamically modulated evidence-integration mechanism". *Plos Computational Biology*, 2016. **12**(2).
[41] Fischenich, A., A. Sekowksi, E. Gorbunov, E. Ponsot, and D. Oberfeld, "A comparison of temporal weights in loudness and brightness judgments". Manuscript in preparation.
[42] Murdock, B.B., "The serial position effect of free recall". *Journal of Experimental Psychology*, 1962. **64**(5): p. 482-488.
[43] Brodie, D.A., and B.B. Murdock, "Effect of presentation time on nominal and functional serial-position curves of free recall". *Journal of Verbal Learning and Verbal Behavior*, 1977. **16**(2): p. 185-200.
[44] Anderson, N.S., and V.D. Burns, "Comparison of presentation rates using a missing item probe test of immediate memory". *Bulletin of the Psychonomic Society*, 1973. **2**(4): p. 200-202.
[45] Waugh, N.C., and D.A. Norman, "Primary memory". *Psychological Review*, 1965. **72**(2): p. 89-104.
[46] Zwicker, E., "Procedure for calculating loudness of temporally variable sounds". *Journal of the Acoustical Society of America*, 1977. **62**(3): p. 675-682.

[47] Chalupper, J., and H. Fastl, "Dynamic loudness model (DLM) for normal and hearing-impaired listeners". *Acta Acustica United with Acustica*, 2002. **88**(3): p. 378-386.

[48] Glasberg, B.R., and B.C.J. Moore, "A model of loudness applicable to time-varying sounds". *Journal of the Audio Engineering Society*, 2002. **50**(5): p. 331-342.

[49] Moore, B.C.J., M. Jervis, L. Harries, and J. Schlittenlacher, "Testing and refining a loudness model for time-varying sounds incorporating binaural inhibition". *Journal of the Acoustical Society of America*, 2018. **143**(3): p. 1504-1513.

[50] Pedersen, B., *Auditory Temporal Resolution and Integration. Stages of Analyzing Time-Varying Sounds*, in *Department of Acoustics, Sound Quality Research Unit*. 2006, Aalborg University: Aalborg, DK.

[51] Hots, J., D. Oberfeld, A. Fischenich, and J.L. Verhey, "Zeitliche Gewichtung der Lautheitswahrnehmung im Lautheitsmodell", *Fortschritte der Akustik - DAGA 2019*. 2019.