Modeling Loudness Enhancement

Daniel Oberfeld
Psychologisches Institut, Universität Mainz, D-55099 Mainz, Germany, Email: oberfeld@uni-mainz.de

Introduction

In the presence of a forward or backward masker, the perceived intensity (loudness) of a brief tone changes as a function of masker level. Consider a loudness matching experiment (masker and target in the first observation interval, comparison tone in the second interval). If masker level is greater than target level, loudness of the target tone is increased [1, 2]. The comparison tone level L_C adjusted by the listener to match target tone loudness will be greater than target level L_T (loudness enhancement: $L_C - L_T > 0$). If masker level is smaller than target level, target tone loudness is reduced (loudness decrement: $L_C - L_T < 0$; [3]).

In the auditory periphery, an intense forward masker causes –if anything– a reduction of the neural response to the target tone (e.g., [4]). Out of this reason, Oberfeld [5] argued that loudness enhancement must be a higher-level effect. In [5], a heuristic model was presented that can account for a broad range of data. The model adopts the "mergence hypothesis" proposed in [3]. It is assumed that the loudness representations of masker and target are merged automatically so that "*[...] the final percept of the target is approximated by a weighted average of the separate sensations each interactor would produce if presented alone*." ([3], p. 606). The second important assumption is that masker loudness receives less weight if masker and target are perceptually different (e.g., in loudness, duration or frequency). Such mechanism has an analog in the well-known effects of target-distractor similarity found in memory experiments.

Model Structure

The model comprises four steps [6].

Step 1: Loudness of Masker and Target

Loudness representations of masker and target are computed according to Zwislocki‘s [7] loudness function. If the masker is a pure tone presented in quiet, its loudness is

$$N_M = K_M\,[(P_M^{\,2} + 2.5\,P_{ThM}^{\,2})^{\theta} - (2.5\,P_{ThM}^{\,2})^{\theta}], \tag{1}$$

where K_M is a scale parameter, P_M is masker pressure and P_{ThM} is the pressure at detection threshold. In the same way, target loudness is modeled as

$$N_T = K_T\,[(P_T^{\,2} + 2.5\,P_{ThT}^{\,2})^{\theta} - (2.5\,P_{ThT}^{\,2})^{\theta}]. \tag{2}$$

A change in threshold (e.g., induced by simultaneous masking or hearing loss) alters the function at low levels, while leaving it relatively unaffected at high levels (loudness recruitment). Sounds differing in, e.g., duration or frequency will produce different loudness values even at high levels. This can be accounted for by choosing different values for the parameters K_M and K_T. The slope parameter θ was found to be 0.3 for a wide range of listening conditions [8].

Step 2: Masker Weight

According to the similarity hypothesis, less mergence will occur if the representations of masker and target differ in one or more dimensions.

The weight assigned to masker loudness is written as

$$p_M = p_{Max} \cdot f(N_M, N_T), \tag{3}$$

where the function $f(N_M, N_T)$ represents the effects of perceptual similarity in the loudness dimension. The effect of similarity in the remaining dimensions is represented by p_{Max}, which is the maximum amount of mergence that will be effective if masker loudness equals target loudness. The function $f(N_M, N_T)$ is chosen in such a way that

$$0 \le f(N_M, N_T) \le 1 \text{ and } f(N, N) = 1. \tag{4}$$

Above that, it is required that $f(N_M, N_T)$ decreases monotonically with the absolute value of the difference between N_M and N_T, approaching 0 at large differences. It also seems reasonable to assume a variant of Weber's law; i.e., masker and target loudness differing by, e.g., 10% result in the same value of p_M independent of target loudness. Given that these conditions are met, the choice of $f(N_T, N_M)$ is certainly arbitrary. The function introduced here can be deducted from a Gaussian discrimination model [6]:

$$f(N_T, N_M) = \left(1 - \mathrm{Erf}\left(\frac{|\log_{10} N_M - \log_{10} N_T|}{\sqrt{2}\sigma}\right)\right), \tag{5}$$

where Erf(x) is the error function. The 'similarity parameter' σ determines how fast p_M decreases with the difference between $\log_{10} N_M$ and $\log_{10} N_T$.

Step 3: Mergence

It is required that the weights p_M and p_T assigned to masker and target loudness, respectively, sum to unity,

$$p_T = 1 - p_M. \tag{6}$$

Therefore, the weighted average between target and masker loudness is predicted to be

$$N_{Merged} = p_M\,N_M + (1 - p_M)\,N_T. \tag{7}$$

Step 4: Loudness Match

The sound pressure level P_C of the comparison tone eliciting a loudness sensation N_C equal to the weighted average N_{Merged} can be found by solving the equation

$$N_{Merged} = K_C\,[(P_C^{\,2} + 2.5\,P_{ThC}^{\,2})^{\theta} - (2.5\,P_{ThC}^{\,2})^{\theta}]. \tag{8}$$

for P_C. The parameters P_T, P_M, P_{ThT}, P_{ThM}, P_{ThC}, K_T, K_M, K_C and θ are known a-priori. Only the parameters p_{Max} and σ need to be estimated when fitting the model. If $K_T = K_M = K_C$, the scale factors cancel out.

Notice, that in performing a loudness match listeners frequently produce a nonzero difference between target and comparison level even if no masker is present [9]. Equally important, $L_C - L_T$ will not necessarily be zero if masker and target are identical (cf. Figure 1, loudness match at $L_M = L_T = 70$ dB SPL). These observations can be accounted for by adding a bias parameter b to Eq. (8),

$$b \cdot N_{Merged} = K_C\,[(P_C^{\,2} + 2.5\,P_{ThC}^{\,2})^{\theta} - (2.5\,P_{ThC}^{\,2})^{\theta}]. \tag{9}$$

The parameter b is assumed to be independent of masker level but is allowed to vary with target level [9].

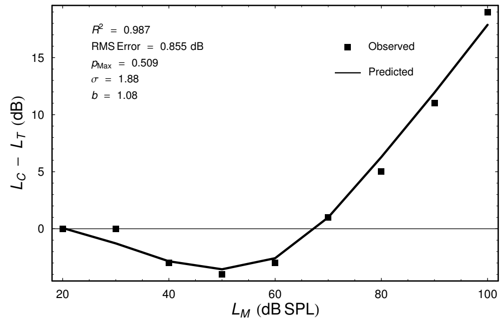

Figure 1: Best fit to data from Exp. 6 reported in [2]. One subject, monaural, 70 dB SPL target. 20-ms, 5-kHz tones. Forward masker level was varied. Fixed parameters: $\theta = 0.3$, $L_{ThM} = L_{ThT} = L_{ThC} = 28$ dB SPL, $K_T = K_M = K_C$. Squares: observed. Line: predicted. The inset shows R^2, RMS error and best fitting parameter values. The bias parameter b was estimated from a match in quiet.

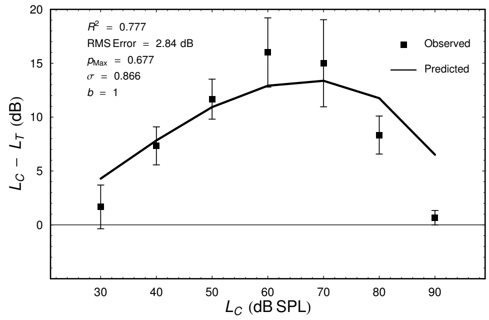

Figure 2: Best fit to mean data (3 subjects) from [10]. Masker level was fixed at 90 dB SPL, comparison tone level was varied and target tone loudness was matched to comparison tone loudness. 1-kHz tones; 100 ms masker, 25 ms target. Fixed parameters: $\theta = 0.3$, L_{ThM} = 11.75 dB SPL, $L_{ThT} = L_{ThC}$ = 17 dB SPL, $K_M = 1.418 \cdot K_T$, $K_T = K_C$ = 1. The bias parameter b was set to 1.0. Error bars show ± 1 SEM.

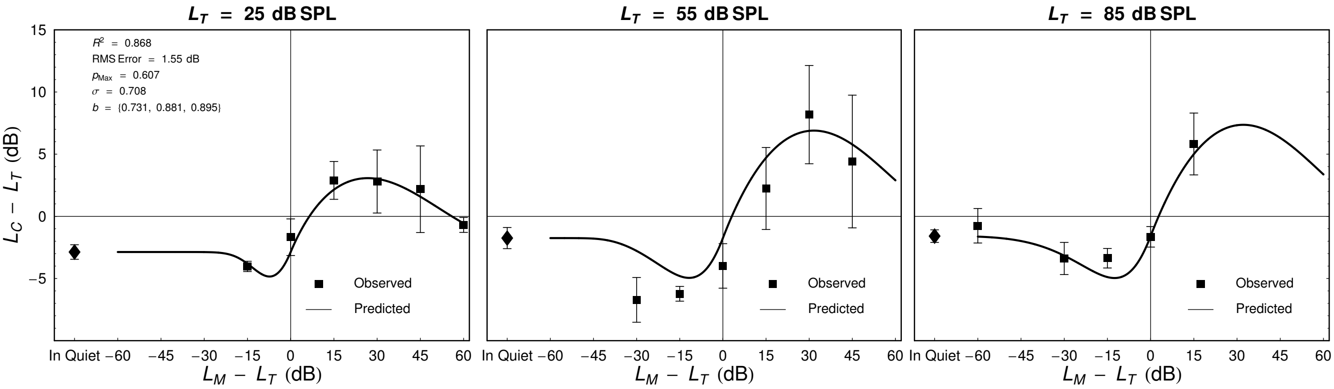

Figure 3: Best fit to mean data (4 subjects) from [5]. 30-ms, 1-kHz tones. $L_M - L_T$ was varied at target level L_T. Fixed parameters: $\theta = 0.3$, $L_{ThM} = L_{ThT} = L_{ThC}$ = 6.41 dB SPL, $K_M = K_T = K_C$. From matches in quiet, a separate value of b was estimated for each target level. Error bars show ± 1 SEM.

Fitting the Model

Figures 1-3 display fits of the model to three data sets [2, 10, 5]. Detection thresholds were reported in each study. In [10], masker duration was longer than target duration, therefore L_{ThM} and K_M were assumed to differ from the respective target tone values (see caption of Figure 2). The bias parameter b was estimated from matches in quiet for the data from [2] and [5] so that it entered the model as a fixed parameter in these cases. For the data reported in [10], b was set to 1.0. Reasonable-to-excellent fits were obtained using two free parameters only (p_{Max} and σ).

To summarize, a simple model was proposed that combines properties of the auditory periphery (loudness function) and higher-level effects (target-distractor similarity). The model correctly predicts the most important empirical observations:

1. Loudness enhancement is observed if masker loudness is greater than target loudness. If masker loudness is smaller than target loudness, loudness decrement is found (Figure 1, [2, 3]).
2. Loudness decrement is less pronounced than loudness enhancement (Figure 1, [2, 3, 5]).
3. Loudness decrement vanishes as masker level approaches detection threshold (Figure 1, [2, 3]).
4. For a fixed target level, the masker-induced loudness change is a non-monotonic function of the masker-target level difference (the mid-difference hump; Figure 3, [5]).
5. Loudness enhancement and decrement are least pronounced at low target levels (Figure 3, [5]). This phenomenon is related to basilar-membrane compression.
6. For a fixed intense masker, loudness enhancement is maximal at intermediate levels (mid-level hump; Figure 2, [10, 11]).

References

[1] Zwislocki, J., & Sokolich, W. (1974). On loudness enhancement of a tone burst by a preceding burst. *Perception and Psychophysics, 16,* 87-90.

[2] Elmasian, R., and Galambos, R. (1975): Loudness enhancement: Monaural, binaural and dichotic. *J. Acoust. Soc. Am., 58*, 229-234.

[3] Elmasian, R., Galambos, R., & Bernheim, A. (1980): Loudness enhancement and decrement in four paradigms. *J. Acoust. Soc. Am., 67*, 601-607.

[4] Bauer, J.W., and Galambos, R. (1975). Evoked potentials in cat auditory nerve: Suppression by prior tonal stimulation. *Perception and Psychophysics, 17*, 43-47.

[5] Oberfeld, D. (2003a). Intensity discrimination and loudness in forward masking: The effect of masker level. In DEGA (Ed.), *Fortschritte der Akustik - DAGA '03.* Oldenburg: DEGA; 606-607.

[6] Oberfeld, D. (2003b): Modeling the effects of forward-masker level on the loudness of a tone. In B. Berglund & E. Borg (Eds.), *Fechner Day 2003. Proceedings of the 19th Annual Meeting of the International Society for Psychophysics.* ISP: Stockholm, Sweden; 211-216.

[7] Zwislocki, J. (1965). Analysis of some auditory characteristics. In R.D. Luce, R.B. Bush, & E. Galanter (Eds.), *Handbook of Mathematical Psychology*, *Vol. III.* New York: Wiley; 3-97.

[8] Hellman, R.P. (1991). Loudness scaling by magnitude scaling: Implications for intensity coding. In S.J. Bolanowski & G.A. Gescheider (Eds.), *Ratio Scaling of Psychological Magnitudes.* Hillsdale: Erlbaum; 215–228.

[9] Hellström, Å. (1985). The time-order error and its relatives: Mirrors of cognitive processes in comparing. *Psychological Bulletin, 97*, 35-61.

[10] Zeng, F.-G. (1994): Loudness growth in forward masking: Relation to intensity discrimination. *J. Acoust. Soc. Am., 96*, 2127-2132.

[11] Plack, C.J. (1996). Loudness enhancement and intensity discrimination under forward and backward masking. *J. Acoust. Soc. Am., 100,* 1024-1030.