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 findings. 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 important additional 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

\[ N_M = K_M \left( (P_M^2 + 2.5 \cdot P_{TM}^2)^{\theta} - (2.5 \cdot P_{TM}^2)^{\theta} \right), \]  

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

\[ N_T = K_T \left( (P_T^2 + 2.5 \cdot P_{TM}^2)^{\theta} - (2.5 \cdot P_{TM}^2)^{\theta} \right). \]

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 \( \theta \) 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

\[ p_M = \frac{P_{Max}}{P_{Mmax}} \cdot f(\theta) \],

where the function \( f(\theta) \) represents the effects of perceptual similarity on the loudness dimension. The effect of similarity on 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 \( f(\theta) \) is chosen in such a way that

\[ 0 \leq f(\theta) \leq 1 \text{ and } f(N, N) = 1. \]

Above that, it is required that \( f(\theta) \) 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 \( f(\theta) \) is certainly arbitrary. The function introduced here can be deduced from a Gaussian discrimination model [6]:

\[ f(N_M, N_T) = \left( 1 - \text{Erf} \left( \frac{\log_{10} N_M - \log_{10} N_T}{\sqrt{2}\sigma} \right) \right). \]

where \( \text{Erf}(x) \) is the error function. The ‘similarity parameter’ \( \sigma \) 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,

\[ p_T = 1 - p_M. \]

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

\[ N_{Merged} = p_M N_M + (1 - p_M) N_T. \]

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

\[ N_{Merged} = K_C \left( (P_C^2 + 2.5 \cdot P_{TM}^2)^{\theta} - (2.5 \cdot P_{TM}^2)^{\theta} \right), \]

for \( P_C \). The parameters \( P_T, P_M, P_{TM}, P_{TM}^2, P_{TM}^3, K_C, K_M, K_T, \theta \) are known a-priori. Only the parameters \( p_{Max} \) and \( \sigma \) 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 \( L_M = L_T = 70 \text{ dB SPL} \)). These observations can be accounted for by adding a bias parameter \( b \) to Eq. (8),

\[ b \cdot N_{Merged} = K_C \left( (P_C^2 + 2.5 \cdot P_{TM}^2)^{\theta} - (2.5 \cdot P_{TM}^2)^{\theta} \right). \]

The parameter \( b \) is assumed to be independent of masker level but is allowed to vary with target level [9].
For a fixed intense masker, loudness enhancement is maximal at intermediate levels (mid-level hump; Figure 2, [10, 11]).

References