Spring 5-31-2005

Tracking the time to recovery after induced loudness reduction (L)

Yoav Arieh
Montclair State University, ariehy@montclair.edu

Karen Kelly
Montclair State University, kellykar@montclair.edu

Lawrence E. Marks
Yale University, lawrence.marks@yale.edu

Tracking the time to recovery after induced loudness reduction (L)

Article *in* The Journal of the Acoustical Society of America · June 2005

DOI: 10.1121/1.1898103

Yoav Arieh
Montclair State University

Lawrence E Marks
The John B. Pierce Laboratory

Tracking the time to recovery after induced loudness reduction (L)

Yoav Arieh[a)]
The Psychology Department, Montclair State University, Montclair, New Jersey 07043

Karen Kelly
The Psychology Department, Montclair State University, Montclair, New Jersey 07043

Lawrence E. Marks
John B. Pierce Laboratory and Yale University, New Haven, Connecticut 06519

(Received 31 December 2004; revised 4 March 2005; accepted 4 March 2005)

In induced loudness reduction (ILR), a strong tone causes the loudness of a subsequently presented weak tone to decrease. The aim of the experiment was to determine the time required for loudness to return to its initial level after ILR. Twenty-four subjects were exposed to 5, 10, 20, or 40 brief bursts of 2500-Hz pure tones at 80-dB SPL (inducers) and then tested in a series of paired comparison trials. Subjects compared the loudness of a weak target (2500 Hz at 60-dB SPL) to the loudness of a comparison tone at 500 Hz previously judged to match the target. The comparison task was repeated until the two tones were again judged equally loud. The results showed that (a) recovery after ILR is a relatively long process with a time scale of minutes, and (b) recovery time increased approximately 20 s with each doubling of the number of inducers. [DOI: 10.1121/1.1898103]

PACS numbers: 43.66.Cb [AJO] Pages: 3381–3384

I. INTRODUCTION

Under appropriate conditions, presenting a relatively strong inducing tone reduces the loudness of a subsequent weaker tone (Arieh and Marks, 2003a; Mapes-Riordan and Yost, 1999; Marks, 1994; Nieder *et al.*, 2003). The phenomenon was initially called recalibration (Marks, 1994), on the premise that it reflects a general principle of intensity processing in the nervous system. Later, the name induced loudness reduction (ILR) was offered to describe the effect specifically in hearing (Scharf *et al.*, 2002). Because this report deals exclusively with loudness, we shall adopt the latter term here.

Four important properties of ILR have been determined since it was first reported (Marks, 1988). First, under optimal conditions the extent of the loudness reduction can reach 10 dB or more (Arieh and Marks, 2003a; Nieder *et al.*, 2003). In other words, a stronger inducer can reduce the loudness, in sones, of the weak target by at least half. Second, ILR is frequency specific, being greatest when the inducer and the target fall within the same critical band (Marks and Warner, 1991). Third, the inducer must precede the target by at least 200 ms (offset–onset) for significant ILR to appear (Arieh and Marks, 2003b). And fourth, ILR is maximal when the SPL of the inducer is 60–80 dB SPL and the inducer and target differ in level by 10–20 dB (Mapes-Riordan and Yost, 1999).

Importantly, ILR does not reflect a response bias but instead is a sensory change, most likely a change in the sensory representation of the intensity of the target tone. Arieh and Marks (2003b) showed that when listeners detect weak tones in a choice decision task, conditions that produce ILR also produce longer response times and higher error rates than do control conditions. The positive relation between response time and error rate is a hallmark of sensory as opposed to decisional change (Luce, 1986).

Still lacking, however, are data about recovery after ILR. After ILR, how long does it take for the loudness of the target tone to return to its original level? The purpose of the current report was to systematically explore this question.

The literature does offer a few clues about recovery after ILR. Arieh and Marks (2003a) showed that 3.3 s after the presentation of the inducer, the loudness of the target tone is still reduced substantially, such that the level of the matching tone is 13 dB less than the matched level without the inducer. Mapes-Riordan and Yost (1998) reported that, even when the interval between the inducer and the target tone is increased to 10 s, ILR remains strong. Marks (1993) noted that ILR largely dissipates when a 60-s pause follows a condition in which ILR is induced. Thus, while ILR arises quickly, within 200 ms, recovery after ILR is a relatively long process requiring some dozens of seconds.

The present study also examines how recovery from ILR depends on the number of inducers. It is reasonable to expect, for instance, that recovery time after ILR will depend on the magnitude of ILR itself: the greater the reduction in loudness, the greater the time to recover. Systematic data about the way magnitude of ILR depends on number of inducers are also lacking. In many studies of ILR, the number of inducers was not controlled nor even reported. For example, in an adaptive procedure often used to study ILR (Arieh and Marks, 2003a; Mapes-Riordan and Yost, 1999), the matching points depend on the listeners' patterns of responses, so the number of trials (equal to the number of

[a)]Electronic mail: ariehy@mail.montclair.edu

presentations of the inducer) can vary considerably between listeners and between experimental conditions.

The effect, if any, of number of inducers bears importantly on the nature of ILR. It is possible, for example, that ILR is an ''all or none'' process that begins anew with each presentation of an inducer. If so, then varying the number of inducers would have no effect on the magnitude of ILR, or on the time for recovery. Alternatively, like so many processes of ''sensory adaptation,'' ILR may reflect the outcome of a suppressive process that builds up through repeated presentation of inducers, and hence requires increasingly greater time to dissipate.

A word is in order as to the logic of the study. Subjects started an experimental session by listening to uninterrupted series of 5, 10, 20, or 40 inducers. Immediately thereafter subjects were presented with a sequence of paired comparison trials. Each trial consisted of the target tone with the same frequency as the inducers and a comparison tone with a different frequency. The latter was previously deemed equal in loudness to the target tone. The subject's task was to judge which tone was louder. We reasoned that if the target tone underwent ILR by the preceding inducers, the listeners would tend to choose the comparison tone as louder. However, as time elapsed ILR would dissipate and loudness would recover. At this stage listeners would increasingly choose the target tone as the louder of the pair. According to this procedure, when the target tone regains its original loudness, listeners' performance would be at or close to chance level. Thus, the point of full recovery was defined as the point where subjects performed the paired comparison task at chance level.

II. RECOVERY FROM ILR AS A FUNCTION OF THE NUMBER OF INDUCERS

A. Method

Twenty women and four men (age range 18–40), all associated with Montclair State University, participated in the study. All reported normal hearing.

Testing sessions were conducted in a sound-treated chamber. A MATLAB program running on a Dell Pentium-IV PC controlled all aspects of stimulus presentation, data collection, and on-line computations. Subjects' responses were entered via the computer keypad. The pure-tone signals were generated by a Tucker-Davis System 3 real time processor at a sampling frequency of 50 kHz. The signals were then appropriately attenuated (Tucker-Davis PA5 module) and delivered binaurally for 50 ms (including 5-ms cosine^2 rise and decay) through calibrated TDH-49 headphones mounted in MX41/AR cushions.

Before the start of each experimental session, the subject performed a baseline task to determine the matching point between the comparison and the target tone. It is critical to establish a reliable matching point because this point serves later to gauge recovery time. In the baseline task, we used a randomized adaptive two-track ascending and descending procedure to determine the level of a 500-Hz tone equal in loudness to the designated target signal—a 2500-Hz tone at 60-dB SPL (details of the procedure can be found in Jesteadt, 1980, and Arieh and Marks, 2003a).

After determining the matching point, the subject served in four experimental conditions divided into two different sessions separated by at least 24 h. Each condition was divided in turn into two continuous parts. In the first part, the subject heard an uninterrupted series of 5, 10, 20, or 40 2500-Hz tones at 80-dB SPL that served as inducers. The 50-ms inducing tones were presented at intervals of 1 s (offset–onset). On the screen, a visual countdown accompanied the inducer sequence and upon its termination a pair of tones was presented for paired comparison. The two tones were the 500-Hz matching level determined at baseline and the 2500-Hz, 60-dB target. The subjects were instructed to judge which tone was louder as quickly as possible without sacrificing accuracy. The order of the two tones was randomized across trials and the tones were presented 1 s apart. The program recorded the subject's response and its latency from the onset of the paired comparison trial.

After the first ten trials the probability p of selecting the target tone as louder was computed. If p fell between 0.4 and 0.6, it was concluded that the subject performed at chance and full recovery had been reached (a final p larger than 0.6 was also considered as complete recovery; however, none occurred in the current study). On the other hand, if p was smaller than 0.4, it was concluded that recovery was not complete and another trial was presented. From this point, the probability of selecting the target tone as louder was continually computed over the ten most recent trials until full recovery had been reached, after which the session terminated.

It was assumed that recovery occurred within the time period that elapsed between the first and the last of the final ten trials. Thus, the midpoint of that time period was taken as the recovery time. For example, if recovery occurred between trial numbers 32 and 41 and these trials occurred 90 and 120 s, respectively, after the start of session, then recovery time was recorded as 105 s. Finally, to minimize the possibility of order effects, we used all 24 possible orders of the four inducer conditions. Each subject received one possible order; the first two conditions were performed in the first session and the last two conditions were performed in the second session.

B. Results

As anticipated, there were no effects of order. The correlation between the recovery time and the serial position of a condition in the presentation order was negligible, $r(94) = -0.039$, $p = 0.78$. The average proportion of selecting the target tone over the first ten trials was 0.229, 0.194, 0.180, and 0.138 and over the last ten trials was 0.421, 0.408, 0.420, and 0.413 for the conditions containing 5, 10, 20, and 40 inducers, respectively. A repeated measures analysis of variance revealed that the final averages did not differ statistically ($F < 1$). Given that $p = 0.5$ represents chance performance, recovery was caught a bit early but equally so in all conditions.

The resolution of our measuring method is the average time for making a paired comparison judgment, which was 3

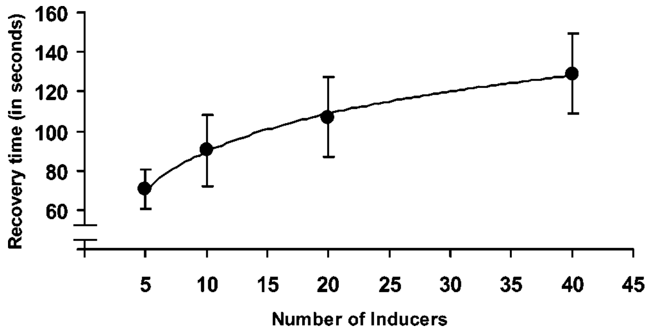

FIG. 1. Averaged recovery times (in seconds) for the 24 listeners plotted against the number of inducers. The error bars represent one standard error of the mean.

s. Thus, the average time to complete the first ten trials that are necessary for the computation of p is 30 s, and the minimal recovery time afforded by our procedure is 15 s. As the data show, these limitations were inconsequential under the stimulus conditions.

The recovery time for each inducer condition was averaged across subjects. The averages are presented in Fig. 1. As can be seen, recovery time increases with the number of inducers. The shortest average recovery time was 70.5 s for the 5-inducer condition and the longest was 129 s for the 40-inducer condition. A repeated-measure analysis of variance showed a main effect of inducer condition, $F(3,69) = 3.64$, $p < 0.05$. A closer look at Fig. 1 also reveals that, within the tested range, the function that relates the number of inducers to recovery time is not linear. In fact, every time the numbers of inducers is doubled an approximately constant amount of time—around 20 s—is added to the recovery time. This suggests that recovery time increases approximately with the logarithm of the number of inducers. The curved line in Fig. 1 shows a log function fitted to the data, computed by regressing recovery time on log number of inducers; the fit is excellent, explaining 99% of the total variance. The small number of data points, however, prohibits a conclusive test of the form of the function. For example, a linear fit explains 93% of the variance in our data set—less than that produced by the logarithmic fit, but still substantial. Thus, at this point any conclusion about the function relating recovery time to number of inducers must be tentative.

One possible limitation of the current procedure is that recovery time is correlated with the number of target presentations. That is, the longer the recovery time the greater the number of comparisons, and, consequently, the greater the number of exposures to the target tones. It is possible that these additional exposures prolong recovery times. To evaluate this possibility we ran ten additional subjects in a control experiment that tested recovery from 40 inducers only, but under two conditions. The first condition replicated the original experiment, in that the paired comparisons began as soon as the induction phase ended. In the second condition, we inserted a 30-s pause between the induction phase and the start of the judgment phase. Thus, in the second condition the subjects were allowed to recover for 30 s before being exposed to the target tones. Consequently, they were exposed, on average, to ten fewer target presentations than the no-pause group. Average recovery times for the no-pause and pause conditions were 113.1 and 119.5 s, respectively, the difference (dependent-pair t-test) not being reliable [$t(9) = 1.2$, ns]. Thus, we conclude that exposure to ten additional target presentations in the no-pause group during recovery had no significant effect on the final recovery time.

III. DISCUSSION

The measurement of recovery after ILR, and indeed the measurement of sensory recovery in general, presents an interesting challenge. Recovery, by definition, is a dynamic process that continuously changes with the passage of time. On the other hand, reliable measurement necessitates several observations of the same event. Observations, especially behavioral observations, however, take time. How can a constantly changing event be measured reliably?

Our solution is to define recovery as a probabilistic process that is computed over a roving window of observations comprising ten successive loudness comparisons. Once the listener can no longer reliably discriminate between the target tone that underwent ILR and a comparison tone that was previously deemed equally loud, recovery has been reached. The method is not perfect, its resolution being about 3 s, and it cannot measure recovery times smaller than 15 s. A crucial consideration is the size of the observation window. Increasing the window's size will increase reliability but will also increase the lower limit of measurement; decreasing the size will decrease the lower limit of measurement but will also decrease reliability. We chose our parameters after careful consideration of the available literature on recovery after ILR. Applying the method to other sensory processes would undoubtedly require adjustments to its parameters.

Our report is the first to measure recovery after ILR to the point where the target's loudness returned to or near to its original level. A main finding is that recovery time increases monotonically with the number of inducer tones presented in the exposure phase. The fastest recovery time, recorded after exposure to five inducers, exceeded 1 min. The slowest recovery time, after exposure to 40 inducers, exceeded 2 min. These results fit previous observations that hinted that recovery after ILR is a long process, on the order of dozens of seconds (Arieh and Marks, 2003a; Mapes-Riordan and Yost, 1998; Marks, 1993). We note here the asymmetry between the onset and the offset of ILR. The onset is a fast process; ILR is already present 200 ms after the presentation of the inducer. The offset has a completely different time scale; dissipation of ILR takes over a minute after exposure to five inducers. Any future model of ILR will have to account for this property.

We also found that a log function provides a good description of the way recovery time depends on the number of inducers. Put simply, adding inducers does not lead to proportional addition of recovery time. In the range used here, each doubling of the number of inducers added another 20 s to recovery time. Regardless of the exact shape of the function that relates the number of inducers to recovery time, the fact that they are positively related is important. Assuming that the magnitude of ILR itself is associated with recovery time, we can surmise that the former is also positively related to the number of inducers. That conclusion supports a model of ILR that has ‘‘memory’’ for previous exposures to inducers: It allows for accumulation of loudness suppression effects over time. On the other hand, the conclusion does not support an ‘‘all or none’’ model in which the suppression effect resets on every exposure to inducer. Admittedly, a direct measurement of the relation between the magnitude of ILR and recovery time will go a long way towards clarifying this issue.

That the magnitude of ILR is related to the number of inducers also has practical implications. Many studies used an adaptive procedure to measure ILR (Arieh and Marks, 2003a, Mapes-Riordan and Yost, 1999). In this procedure the inducer is presented on each trial but the number of trials depends on the pattern of responses and thus often varies between listeners and between experimental conditions. The fact that the number of inducers has not been controlled in many earlier studies might contribute to the significant variability that is often observed in studies of ILR.

ACKNOWLEDGMENTS

Preparation of this article was supported in part by Grant DC00271-17 from the National Institutes of Deafness and Other Communication Disorders to L.E.M. We would like to thank Scott Parker, Christopher Plack, and Andrew Oxenham for their cogent comments on the earlier version of this paper.

Arieh, Y., and Marks, L. E. (**2003a**). ‘‘Time course of loudness recalibration: Implications for loudness enhancement,’’ J. Acoust. Soc. Am. **114**, 1550–1556.

Arieh, Y., and Marks, L. E. (**2003b**). ‘‘Recalibrating the auditory system: A speed-accuracy analysis of intensity perception,’’ J. Exp. Psychol. Hum. Percept. Perform. **29**, 523–536.

Jesteadt, W. (**1980**). ‘‘An adaptive procedure for subjective judgments,’’ Percept. Psychophys. **28**, 85–88.

Luce, R. D. (**1986**). *Response Times* (Oxford University Press, New York).

Mapes-Riordan, D., and Yost, W. A. (**1998**). ‘‘Temporal properties of loudness recalibration,’’ Proceedings, 16th Int. Cong. Acoust., pp. 2361–2362.

Mapes-Riordan, D., and Yost, W. A. (**1999**). ‘‘Loudness recalibration as a function of level,’’ J. Acoust. Soc. Am. **106**, 3506–3511.

Marks, L. E. (**1988**). ‘‘Magnitude estimation and sensory matching,’’ Percept. Psychophys. **43**, 511–525.

Marks, L. E. (**1993**). ‘‘Contextual processing of multidimensional and unidimensional auditory stimuli,’’ J. Exp. Psychol. Hum. Percept. Perform. **19**, 227–249.

Marks, L. E. (**1994**). ‘‘Recalibrating the auditory system: The perception of loudness,’’ J. Exp. Psychol. Hum. Percept. Perform. **20**, 382–396.

Marks, L. E., and Warner, E. (**1991**). ‘‘Slippery context effect and critical bands,’’ J. Exp. Psychol. Hum. Percept. Perform. **17**, 986–996.

Nieder, B., Buus, S., Florentine, M., and Scharf, B. (**2003**). ‘‘Interactions between test-and inducer-tone durations in induced loudness reduction,’’ J. Acoust. Soc. Am. **114**, 2846–2855.

Scharf, B., Buus, S., and Nieder, B. (**2002**). ‘‘Loudness enhancement: induced loudness reduction in disguise? (L),’’ J. Acoust. Soc. Am. **112**, 807–810.