

# The role of temporal-masking patterns in the determination of subjective duration and loudness for ramped and damped sounds

Dennis T. Ries<sup>a)</sup>

*School of Hearing, Speech and Language Sciences, Ohio University, Athens, Ohio 45701*

Robert S. Schlauch

*University of Minnesota, Speech-Language-Hearing Sciences, Minneapolis, Minnesota 55455*

Jeffrey J. DiGiovanni

*School of Hearing, Speech and Language Sciences, Ohio University, Athens, Ohio 45701*

(Received 15 August 2007; revised 17 September 2008; accepted 22 September 2008)

The association between temporal-masking patterns, duration, and loudness for broadband noises with ramped and damped envelopes was examined. Duration and loudness matches between the ramped and damped sounds differed significantly. Listeners perceived the ramped stimuli to be longer and louder than the damped stimuli, but the outcome was biased by the stimulus context. Next, temporal-masking patterns were measured for ramped- and damped-broadband noises using three (0.5, 1.5, and 4.0 kHz) 10 ms probe tones presented individually at various temporal delays. Predictions of subjective duration derived from masking results underpredicted the matching results. Loudness estimates derived from models that assume persistence of neural activity after stimulus offset [Glasberg B. R., and Moore, B. C. J. (2002). "A model of loudness applicable to time-varying sounds," *J. Audio. Eng. Soc.* **50**, 331–341; Chalupper, J., and Fastl, H. (2002) "Dynamic loudness model (DLM) for normal and hearing-impaired listeners," *Acust. Acta Acust.* **88**, 378–386] were greater for ramped sounds than for damped sounds and were close to the average results obtained via the matching task. Differences in simultaneous-masked thresholds for these stimuli could not account for the loudness-matching results. Decay suppression of the later occurring portion of the damped stimulus may account for the differences in perception due to the stimulus context; however, a parsimonious implementation of this process that accounts for both subjective duration and loudness judgments remains unclear. © 2008 Acoustical Society of America.

[DOI: 10.1121/1.2999342]

PACS number(s): 43.66.Ba, 43.66.Cb, 43.66.Dc, 43.66.Lj [BCM]

Pages: 3772–3783

## I. INTRODUCTION

Ramped (rising intensity) and damped (falling intensity) stimuli have been employed to examine how the auditory system processes sounds that have identical long-term spectra yet differ in their temporal characteristics. Prior studies have reported results for a variety of auditory phenomena including intensity discrimination (Schlauch *et al.*, 1998), subjective duration (DiGiovanni and Schlauch, 2007; Grassi and Darwin, 2006; Schlauch *et al.*, 2001a, 2001b), loudness (Irino and Patterson, 1996; Neuhoff 1998, 2001; Maier *et al.*, 2004; Stecker and Hafter, 2000), and tonality (Patterson, 1994a, 1994b). All of these studies, except that on intensity discrimination (Schlauch *et al.*, 1998), report perceptual differences between ramped and damped stimuli of the same physical duration. Possible explanations for this perceptual asymmetry have included perceptual constancy (DiGiovanni and Schlauch, 2007; Grassi and Darwin, 2006; Stecker and Hafter, 2000) and persistence of excitation within the central auditory system (DiGiovanni and Schlauch 2007).

Stecker and Hafter (2000) proposed that a cognitive effect is responsible for at least a portion of the perceptual differences that result from manipulations of the temporal envelope of a sound. They proposed that perceptual constancy caused sounds with damped envelopes to be perceived as softer than sounds with ramped envelopes. Perceptual constancy is a cognitive phenomenon that describes how an individual may obtain similar perceptual experiences for a given stimulus across a variety of environmental settings. Visual size constancy is a well known example of this phenomenon. In the visual realm, when objects move farther from an observer, they are not judged as smaller, provided that cues for distance are available, as the brain compensates for distance to maintain size constancy. In a similar vein, Stecker and Hafter (2000) proposed that the brain performs a constancy-mode interpretation of sounds with damped envelopes. According to their hypothesis, sounds with damped envelopes are parsed into two pieces. The earliest or first part represents a sound source, and the second part represents reverberation. The part judged as reverberation is discarded when making judgments about source loudness and duration. This reduces the loudness and perceived duration of sounds with damped envelopes. By contrast, sounds with ramped

<sup>a)</sup>Electronic mail: ries@ohio.edu

envelopes do not fit the pattern of a sound source and its echo, so they are judged as a single source without reverberation.

An alternative explanation is conceivable, however, as auditory sensation in response to a sound does not end abruptly upon termination of the stimulus but has been suggested to decay in an exponential manner for some time afterward [e.g., see Miller (1948); Plomp (1964), and Zwillocki *et al.* (1959)]. Differences in persistence of excitation within the central auditory system that could contribute to perceptual differences between ramped and damped sounds would be expected, as the former ends at a much higher level than the latter. This would lead to a greater amount of forward masking for the ramped than for the damped sound [e.g., see Fastl (1976, 1976/1977, 1979)]. The forward-masking patterns could be interpreted as an estimate of the degree to which the excitation produced by these two sounds extends beyond the stimulus offset [e.g., see Fastl (1984) and Fastl and Zwicker (2007)]. These patterns purportedly represent neural excitation at some unspecified level in the auditory system, and differences between these patterns would be expected to produce differences in perception if persistence of excitation affects perception. Zwicker (1977) as well as Fastl (1977) and Fastl *et al.* (2002) modeled loudness and subjective duration from these patterns. In their modeling, they assumed that some portion of the neural excitation that continued beyond the offset of an auditory stimulus contributed to the loudness and subjective duration of the sound and that the extent of this effect could be estimated from forward-masking data. Results from temporal-masking studies reveal that forward masking has a much longer time course for decay than the buildup of backward masking [e.g., see Fastl (1976, 1976/1977, 1977)]. Hence, estimates of the nonsimultaneous excitation attributed to a stimulus derived from forward-masking data would have a greater influence on modeling of loudness and subjective duration than would estimates derived from backward-masking data. Given this expected asymmetry in the masking patterns, a sound with a ramped envelope that ends at a high level will have its excitation extended by persistence of excitation far beyond that of a sound with a damped envelope that ends at a low level. This is consistent with sounds with ramped envelopes being perceived as louder and longer than sounds with damped envelopes, even though they have the same physical duration.

The present study measures temporal-masking patterns for ramped- and damped-broadband-noise maskers and compares them to loudness and subjective duration estimates. The loudness and duration measures were obtained via a matching procedure, whereas the temporal-masking patterns were measured using a short-duration probe placed at various temporal locations with respect to the ramped and damped sounds. The primary purpose of this study was to determine the extent to which excitation patterns derived from masking patterns can account for observed differences in loudness and perceived duration for these sounds. These measures were intended to provide for a better understanding of the mechanisms responsible for the perceptual differences between ramped and damped sounds.

## II. EXPERIMENT 1: DURATION MATCHING

Duration matches were obtained for sounds as a function of duration (10–500 ms). In the control condition, subjects adjusted the duration of a broadband noise with a rectangular envelope to match that of a broadband noise with the same envelope. In the two experimental conditions, listeners adjusted the duration of broadband noise with a damped envelope until it was perceived to have the same duration as a standard stimulus with a ramped envelope (experimental condition 1) or vice versa (experimental condition 2).

### A. Stimuli

A custom-designed 16 bit digital-to-analog converter generated the stimuli. Broadband noise was generated at a sampling rate of 26.0 kHz with anti-aliasing filters set to 10.0 kHz. The sounds had a rectangular, a ramped, or a damped envelope. The damped envelope was generated by an exponential decay with a time constant set to 1/5th of the signal duration and is given by

$$x(t) = e^{(-5t/T)}w(t), \quad (1)$$

where  $w$  is the rectangular-gated noise,  $t$  is time in seconds, and  $T$  is the duration of  $w$  in seconds. The ramped-exponential rise was produced by reversing the order of the points in the array generated for the damped sound prior to playback. The value of  $T$  for the fixed-duration standard was 10, 25, 50, 100, 200, or 500 ms. The value of  $T$  for the comparison stimulus was adjustable through the use of a slide mechanism. As the range of adjustment was determined randomly for each run, the duration value associated with the matched position from the preceding run was rarely, if ever, the same as that of a subsequent run. The range of adjustment available was always much broader than the range of the subjects' matches. The standard and comparison were presented alternately with a 500 ms interval between the end of the standard and the beginning of the comparison and a 1000 ms delay between the end of the comparison and the beginning of the subsequent standard stimulus. The rms level of the rectangular-gated noise, without the exponential decay or rise, was 70 dB SPL, corresponding to a spectrum level of 30 dB.

### B. Subjects

Five adults (one male and four females) aged 18–30 years participated in the matching experiment. All of the subjects had normal hearing (thresholds  $\leq 15$  dB HL) at octave frequencies from 0.25 to 8.0 kHz and had no prior experience listening in matching experiments.

### C. Procedure

Subjects were presented with a fixed-duration standard stimulus followed by an adjustable comparison. Subjects were instructed first to bracket the point of perceived duration equality in order to establish the positions along the slider range that produced comparison stimuli that were clearly longer and shorter than the standard. After this was accomplished, they were asked to adjust the slider to the

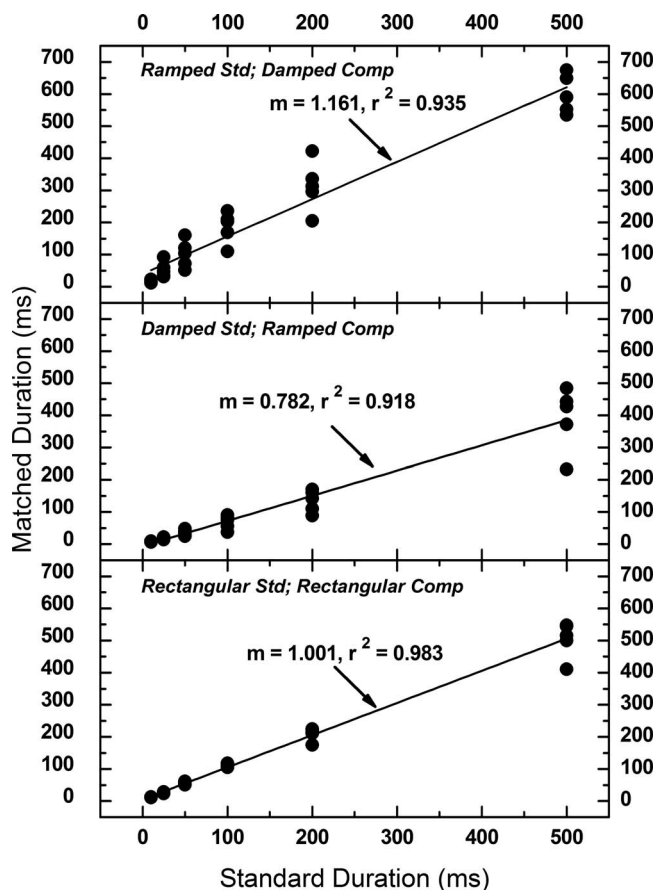


FIG. 1. Scatter plot of the duration-matching results in ms (y-axis) for the ramped standard, damped comparison (upper panel), damped standard, ramped comparison (middle panel), and control conditions (lower panel) as a function of standard duration (x-axis) for all subjects participating in experiment 1. The lines represent least-squares linear fits to the data. The proportion of the variance that can be accounted for by each linear regression analysis is given by the value of  $r^2$ .

position that produced standard and comparison sounds with the same perceived duration. Once a matching duration was obtained, the subjects pushed a button on a response box and the matching duration was stored.

Each subject completed the matching task ten times for each standard-comparison pair (damped standard versus ramped comparison, ramped standard versus damped comparison, and rectangular standard versus rectangular comparison) for each of the six standard durations. Stimuli were presented to the right ear of each subject via an insert earphone (ER-3A, Etymotic Research). Subjects listened individually in a double-walled sound-treated booth (IAC Model 1204-A, Industrial Acoustics Corp.).

#### D. Results

Figure 1 illustrates average duration matches (y-axis) as a function of the standard duration (x-axis) for the control condition (lower panel) and the experimental conditions (middle and upper panels). Linear regression and correlation analyses reveal a strong association between the duration of the standard and the duration of the matched comparison ( $r^2$  values  $>0.91$  for all correlations). Furthermore, the slope of the regression line fitted to each data set differed across con-

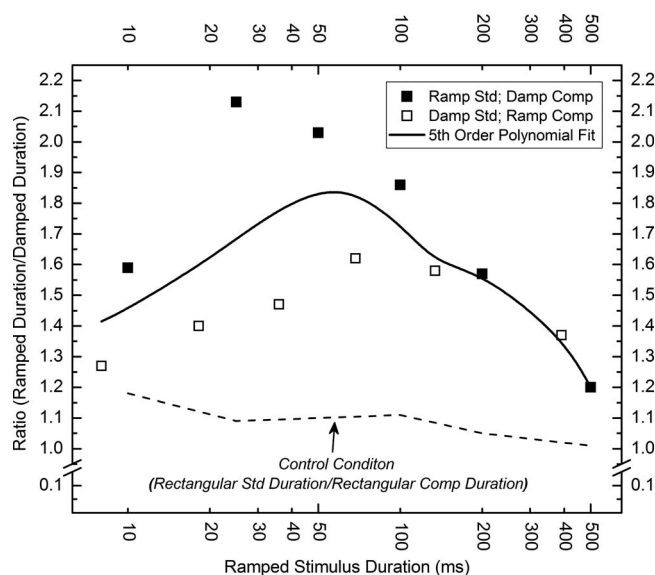


FIG. 2. Mean ramped/damped duration ratios (y-axis) as a function of ramped-stimulus duration (x-axis) as measured using the matching task. The results for the ramped-standard, damped-comparison condition are depicted by filled squares, and the results for the damped-standard, ramped-comparison condition are depicted by open squares. The solid line represents a fifth-order polynomial fit to the data. The data from the control condition (ratio of the durations of the matched rectangular comparison/rectangular standard) are represented by the dashed line.

ditions, indicating that the duration-matching behavior was affected by the envelope of the standard and comparison pair. Specifically, the slope was greater than 1.0 for the match of the damped comparison to the ramped standard (upper panel), whereas the opposite (slope  $<1.0$ ) occurred for the ramped comparison and damped standard condition (middle panel). The slope of the line in the control condition (lower panel) was 1, indicating no time-order bias.

Figure 2 shows the results expressed as the ratio of the duration of the ramped sound to that of the damped sound at equal perceived duration when the damped sound served as the standard (open squares) and when the ramped sound served as the standard (filled squares). The ratio of the comparison to standard duration (dashed line) for the control condition is provided for reference. The ratios at each duration, as well as a fifth-order polynomial fit (solid line) to the ratios for the experimental conditions, illustrate that the ramped stimuli were perceived to be longer than the damped stimuli across all durations. Furthermore, the disparity in perceived duration between the ramped and damped sounds varies as a function of duration with a maximal difference at around 50–100 ms on the fitted function. The ratios for the experimental conditions were greater than those for the control condition for all durations.

The matched durations across subjects were analyzed using a repeated-measure analysis of variance (ANOVA) with standard duration and condition (standard and comparison envelope pair) as factors. There was a significant effect of standard duration [ $F(5, 20)=482.70$ ,  $p < 0.0001$ , with the Geisser–Greenhouse adjustment] and condition [ $F(3, 8)=17.28$ ,  $p=0.014$ ]. As expected, there was a significant interaction between condition and duration [ $F(10, 40)=6.28$ ,  $p=0.035$ ], indicating that the matched duration for a given

TABLE I. Results of the *post hoc* analyses for the duration-matching data comparing stimulus envelope condition and duration. An asterisk indicates that the results of the analyses were significant at an alpha level of 0.05. NS indicates that no significant difference was found for a given comparison.

Condition comparison	Duration (ms)						
	Overall	10	25	50	100	200	500
Control vs Damped standard/ramped comparison	*	NS	NS	NS	NS	*	*
Control vs Ramped standard/damped comparison	*	NS	NS	*	*	*	*
Damped standard/ramped comparison vs Ramped standard/damped comparison	*	NS	NS	*	*	*	*

standard-comparison pair was dependent on both the temporal envelope of the stimuli and the duration of the standard.

As the *F*-statistic was significant for condition and duration, *post hoc* comparisons for these factors and their interaction were evaluated using Fisher's protected least significant difference (LSD) multiple comparison test (Ryan, 1959) with an alpha level of 0.05. The results of the *post hoc* analyses are provided in Table I. The results combined across durations indicated that all conditions differed significantly. The duration matches obtained in the damped standard/ramped comparison and ramped standard/damped comparison conditions differed significantly from each other for standard durations of 50 ms and greater. When compared to the control condition, however, it is clear that the differences were not symmetrical. That is, the ramped standard/damped comparison results differed from those for the control condition for more of the standard durations than did the results obtained for the damped standard/ramped comparison condition. Similar asymmetrical outcomes (i.e., order/context effects) for duration have been obtained by others (Grassi, personal communication). Not surprisingly, the *post hoc* results for the duration for each standard duration combined across envelope type, provided in Table II, indicated that the matches obtained for a given standard duration differed significantly from those obtained at all other durations, except when the results obtained using the 10 and 25 ms standards were compared to each other.

### E. Discussion

These results are consistent with those reported previously (DiGiovanni and Schlauch, 2007; Grassi and Darwin, 2006; Schlauch *et al.*, 2001b). That is, damped sounds are

TABLE II. Results of the *post hoc* analyses for the duration-matching data at each standard duration combined across envelope type. See Table I for the description of abbreviations and symbols.

Standard duration (ms)	Comparison duration (ms)					
	10	25	50	100	200	500
10	...	NS	*	*	*	*
25		...	*	*	*	*
50			...	*	*	*
100				...	*	*
200					...	*
500						...

perceived as being shorter than ramped sounds of equivalent duration. This difference is related to a disparity in perception and is not a function of the measurement method, as subjects were able to match the durations of two sounds with rectangular envelopes accurately (see Fig. 1). Unlike the results of Schlauch *et al.* (2001b), however, the present results show a greater difference in the ramped/damped matched duration when the ramped sound was the standard than when the damped sound was the standard (see Fig. 2). The ramped/damped duration ratios for the mid-duration stimuli (i.e., 25, 50, and 100 ms) when the ramped sound served as the standard are larger than those reported earlier (compare to the lower left panel of Fig. 6 in Schlauch *et al.*, 2001b). The matches reported for the damped standard condition are comparable between the two studies. Interestingly, the line fit to the ratios in Fig. 2 more closely follows the trend exhibited by the magnitude estimation data of Schlauch *et al.* (upper left panel of Fig. 6 in Schlauch *et al.*, 2001b) than it follows their matching data. In addition, the results appear to be consistent with the results of Grassi and Darwin (2006) in that the difference in the perceived duration between ramped and damped stimuli of equivalent durations decreases for longer-duration stimuli (e.g., 200–500 ms in this study and 250–1000 ms in Grassi and Darwin's (2006) study).

### III. EXPERIMENT 2: LOUDNESS MATCHES

The stimuli used in this experiment were the same as those employed in experiment 1. Loudness matches were obtained for two experimental conditions and one control condition. In the control condition, subjects adjusted the level of a broadband noise with a rectangular envelope to match the loudness of a broadband noise with the same envelope. In the two experimental conditions, listeners adjusted the level of broadband noise with a damped envelope until it was perceived to have the same loudness as the standard stimulus with a ramped envelope (condition 1) or vice versa (condition 2).

#### A. Subjects

Four adults (one male and three females) aged 18–30 years participated and were selected using the same criteria as for experiment 1. Three of the four subjects also participated in experiment 1. The new subject had no prior experience listening in matching experiments.

## B. Procedure

The matching task required subjects to adjust a slide mechanism on a response box to equate the loudness of an adjustable comparison and a fixed-level standard stimulus. The initial level of the comparison stimulus was determined from the position of the slide at the beginning of each run. The levels associated with any slide position varied randomly between runs. The maximum range of adjustment of 38 dB was determined based on pilot data and always included the 70 dB SPL standard within this limit. The position of the 70 dB SPL point within this range was offset randomly above and below the center position of the slider on any given run by as much as  $\pm 11$  dB, ensuring that subjects could not simply set the slide to the same place each time. In addition, this range was set so that there would be a minimum of 8 dB of adjustment possible on either side of the standard level. A 95 dB SPL ceiling was adopted to ensure that the upper end of the range did not become uncomfortably loud. In cases where the offset of the standard from the midpoint of the slider range was more than +6 dB, the total range was reduced on the upper end to limit this value to 95 dB SPL. When the offset of the standard from the midpoint of the slider range was less than -6 dB (i.e., -7 to -12 dB), the lower end of the range was set to 45 dB SPL to preserve the symmetry within the adjustment range adopted for positive offsets of more than 6 dB. With the offset and the maximum and minimum levels set as specified, the extreme lower and upper values of the adjustment range were 45–78 dB SPL for an offset of -11 dB and 62–95 dB SPL for an offset of +11 dB, respectively. The slight differences in the adjustment range across the various offset values were expected to have limited impact on the loudness judgments (Gabriel *et al.*, 1997). Asymmetrical positioning of the standard within the adjustment range, however, can produce disparate loudness judgments for a given level (Gabriel *et al.*, 1997). To lessen this effect, the offset value of the standard within the adjustment range was selected randomly for each matching run completed by a subject, and the results were averaged over ten such runs for each subject. That is, we expected that any bias, if present, when averaged across runs and subjects, would be mitigated such that it would have a negligible impact on the overall results.

Subjects were presented with a fixed-intensity standard stimulus followed by an adjustable-intensity comparison stimulus. Standard and comparison stimuli were presented alternately with a 500 ms delay between the end of the standard and the beginning of the comparison and a delay of approximately 1000 ms between the end of the comparison and the beginning of the next standard presentation. When the subject moved the slide, the intensity of the comparison was altered during the next 1000 ms interstimulus interval, and the effect of this adjustment occurred in the next comparison presentation. Subjects were instructed to bracket the point of perceived-loudness equality and to push a button when they felt that the two sounds were equally loud. That is, they were instructed to be sure to adjust the slider until the

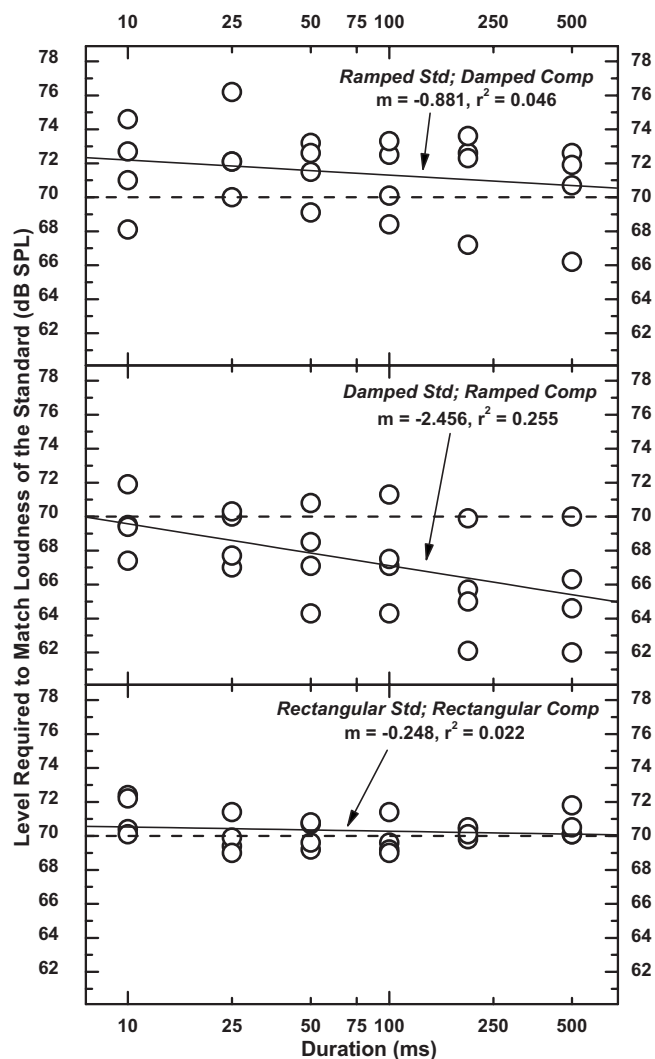


FIG. 3. Similar to Fig. 1, except that the results are for loudness matching.

sound was easily judged to be louder and easily judged to be softer, and then to refine their judgments between those two points.

## C. Results

Figure 3 illustrates average loudness matches (y-axis) as a function of the duration (x-axis) for the control condition (lower panel) and the experimental conditions (middle and upper panels). Linear regression and correlation analyses revealed that the level (dB SPL) required to match the loudness of the comparison to that of the standard was independent of stimulus duration for the control (slope =  $-0.248$  dB/log(ms),  $r^2=0.022$ ,  $p=0.492$ , see the bottom panel) and ramped standard/damped comparison (slope =  $-0.881$  dB/log(ms),  $r^2=0.046$ ,  $p=0.315$ , see the upper panel) conditions, whereas there was a significant correlation between these parameters for the damped standard/ramped comparison (slope =  $-2.456$  dB/log(ms),  $r^2=0.255$ ,  $p=0.012$ , see middle the panel) condition. For the ramped standard/damped comparison condition, the level difference required to equate loudness decreased slightly as the duration of the standard increased, whereas the difference increased as the standard duration increased for the damped standard-

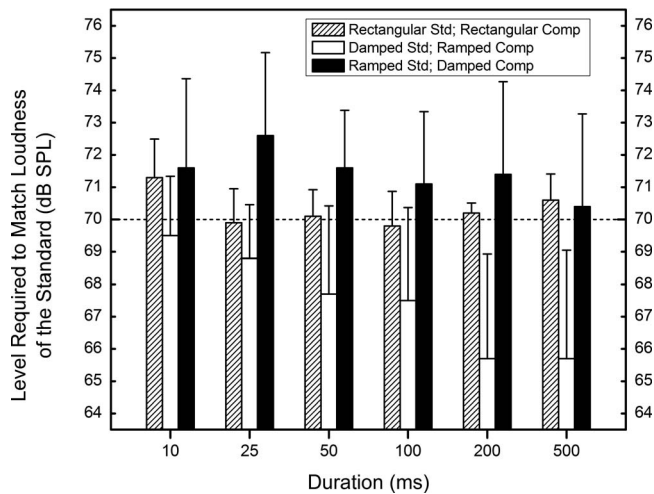


FIG. 4. Mean loudness-matching results (y-axis) for ramped and damped stimuli and for the control condition at each standard duration (x-axis). Black bars depict the results for the ramped-standard and damped-comparison condition, white bars depict the results for the damped-standard and ramped-comparison condition, and hatched bars depict the results for the control condition. The dashed line represents the level of the standard for each condition and duration.

ramped comparison condition. Furthermore, the slope of the regression line and the intercept on the y-axis differed across conditions, indicating that the loudness-matching behavior was affected by the envelope of the standard and comparison pair. Specifically, the level required to match the loudness of a damped comparison to a ramped standard (upper panel) decreased progressively as the duration of the standard increased [slope =  $-0.881$  dB/log(ms)] starting approximately at 2.2 dB above the level of the standard at 10 ms and approaching the level of the standard by 500 ms. In contrast, the level required to match the loudness of the ramped comparison to the damped standard was approximately 0.4 dB below that of the standard at 10 ms and progressively decreased [slope =  $-2.456$  dB/log(ms)] to more than 4.5 dB as the duration of the standard increased to 500 ms. These effects appear to be related to the differences in the temporal envelopes of the ramped and damped sounds as the slope of the line in the control condition (lower panel) was zero with an intercept (70.8 dB SPL) close to the level of the standard (70 dB SPL).

Figure 4 shows the average loudness matches for the control (hatched bar), ramped standard/damped comparison (dark bar), and damped standard/ramped comparison (white bar) conditions for each standard duration. The level of the standard in each condition is denoted by the dashed line. The results indicate that (1) subjects could perform that task consistently and accurately as indicated by the results for the control condition; (2) loudness matching at the shortest duration was slightly less accurate (see control condition results at 10 ms); (3) the level difference for equal loudness between ramped standard and damped comparison conditions did not differ markedly across duration, although the level differences for equal loudness tended to be smaller as duration increased; and (4) the level difference for equal loudness with damped comparison and ramped standard conditions increased over the first 200–500 ms. The latter two points

indicate that the levels of the comparisons required for equal loudness are not symmetrical about the level of the standard across the duration for the two experimental conditions. That is, a difference in the matched level between ramped and damped sounds occurred when the damped sound served as the standard and the ramped sound was the comparison, but not when the situation was reversed.

The loudness matches were analyzed further using a repeated-measure ANOVA with stimulus duration and condition (standard and comparison envelope pair) as factors. The Geisser–Greenhouse correction was applied. There was a significant effect of stimulus duration [ $F(5, 15) = 7.54$ ,  $p = 0.0375$ ] and condition [ $F(2, 6) = 9.58$ ,  $p = 0.0231$ ]. The interaction between condition and duration was insignificant [ $F(10, 30) = 1.18$ ,  $p = 0.369$ ]. A *post hoc* comparison of the outcomes for condition indicated that the level of the matched comparison differed significantly (Fisher's protected LSD,  $p < 0.05$ ) between the damped-standard, ramped-comparison condition and the ramped-standard, damped comparison condition as well as the damped-standard, ramped comparison condition and the control condition, but the results for the ramped-standard, damped comparison condition did not differ significantly from those for the control condition.

#### D. Discussion

The results are in accordance with previous loudness measures (Stecker and Hafter, 2000) involving ramped and damped stimuli. The level required to match the loudness of a ramped or damped standard to that of a damped or ramped comparison is dependent on which sound serves as the standard (see Fig. 4). That is, the level difference needed to match loudness is greatest when the damped sound serves as the standard and increases with increases in stimulus duration up to roughly 200 ms. In contrast, the extent of the level differences encompasses a smaller range overall and does not change significantly with increases in duration when the ramped sound serves as the standard. The mean of the level difference required to match loudness across subjects at each duration was about 1.0–2.0 dB across duration. In contrast, the mean of the level difference required to match loudness across subjects at each duration when the damped sound served as the standard increased from approximately 0.5 dB at 10 ms to over 4.0 dB at 200 ms. The mean matched level at 500 ms was similar to that at 200 ms. This same asymmetry in the size of the ramped-damped loudness difference was reported by Stecker and Hafter (2000). The auditory system appears to be biased toward perceiving a difference in loudness between two sounds when a sound with a damped envelope precedes other stimuli (e.g., sounds with ramped or rectangular envelopes) of similar duration.

#### IV. EXPERIMENT 3: TEMPORAL-MASKING PATTERNS FOR A BROADBAND-NOISE MASKER

Ramped- and damped-broadband-noise stimuli and a 10 ms probe tone were used to obtain temporal-masking patterns for three masker durations (10, 50, and 500 ms). The durations of the maskers correspond to the two end points of

the duration range used in experiments 1 and 2 and a mid-duration value (close to the geometric mean of the durations at the end points).

### A. Stimuli

The 10, 50, and 500 ms noise maskers used in this experiment were identical to the equivalent duration ramped- and damped-broadband-noise stimuli used in the matching experiments. Thresholds in quiet and in the presence of the broadband maskers were obtained for a 10 ms probe tone presented before, during, and after the masker. The tones were gated on and off with 5 ms minimum-three-term Blackman–Harris-shaped (Nuttall, 1981) rise and fall ramps and had no steady state. The peak in the envelope of the probe occurred 5 ms after the onset of the probe tone. For the 10 ms masker, the temporal placement of the peak amplitude of the probe tone with respect to the onset of the masker occurred at  $-5$ ,  $5$ ,  $15$ , and  $30$  ms for the ramped masker and at  $-5$ ,  $5$ ,  $15$ , and  $25$  ms for the damped masker. For the 50 ms masker, the corresponding times were  $-15$ ,  $-5$ ,  $5$ ,  $15$ ,  $25$ ,  $35$ ,  $45$ ,  $65$ ,  $75$ , and  $85$  ms for the ramped masker and  $-15$ ,  $-5$ ,  $5$ ,  $15$ ,  $25$ ,  $35$ ,  $45$ ,  $55$ ,  $65$ , and  $75$  ms for the damped masker. For the 500 ms masker condition, the corresponding times were  $-15$ ,  $-5$ ,  $5$ ,  $125$ ,  $250$ ,  $375$ ,  $415$ ,  $455$ ,  $495$ ,  $515$ ,  $530$ ,  $545$ ,  $575$ , and  $645$  ms for the ramped masker and  $-75$ ,  $-35$ ,  $-15$ ,  $-5$ ,  $5$ ,  $45$ ,  $85$ ,  $125$ ,  $250$ ,  $375$ ,  $495$ ,  $505$ , and  $515$  ms for the damped masker. Probe frequencies were 0.5, 1.5, and 4.0 kHz.

### B. Subjects

Three adult subjects (two males and one female) aged 28–30 years were selected according to the criteria employed in experiments 1 and 2. Only one of the subjects, the first author, had prior experience listening in temporal-masking experiments. The remaining two subjects were given practice on the task in quiet until they felt comfortable with the procedure and produced consistent thresholds (approximately five to seven runs). They were also given 1–2 h of practice listening in various backward-, simultaneous-, and forward-masked conditions.

### C. Procedure

Thresholds in quiet and at various probe-signal delays were obtained using a three-interval forced-choice task that targeted 70.7% of the psychometric function (Levitt, 1971). Lights marked the observation intervals, and 500 ms silent periods separated each interval within a trial. Subjects were instructed to pick the interval that sounded different. Signals decreased in level following two correct responses and increased in level following one incorrect response. The step size was 3 dB. The first three reversals were discarded, and the threshold for each 80-trial run was based on the average of the levels at which the remaining reversals occurred. A run was discarded if fewer than six reversals were obtained; a typical number of reversals was 14.

The first run began with the signal level about 10 dB above a threshold value estimated from the pilot data. Subsequent runs began with the signal level about 10 dB above

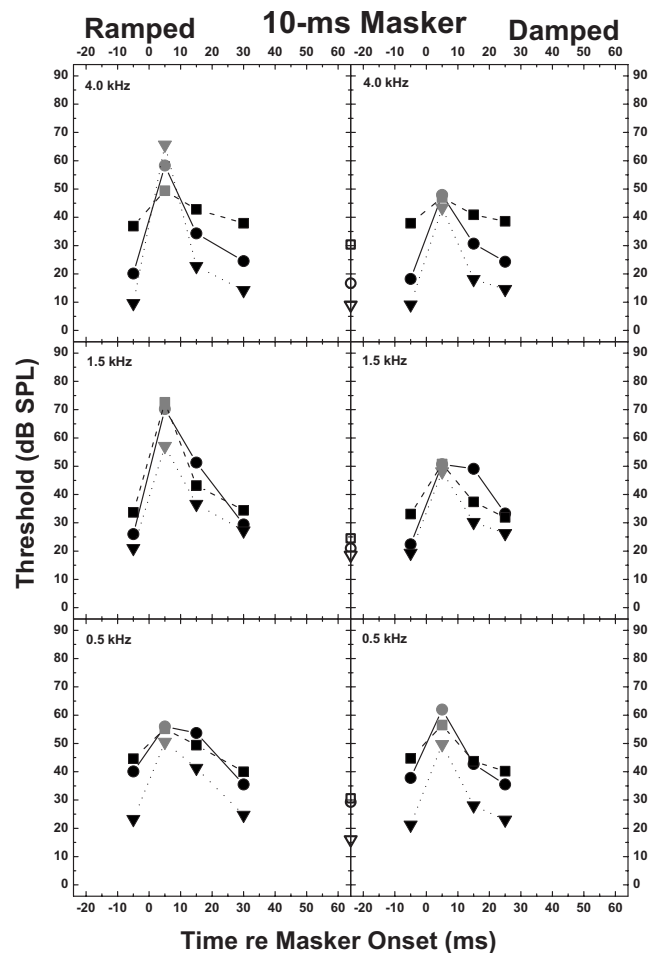


FIG. 5. Temporal-masking patterns for 10 ms ramped (left panels) and damped (right panels) broadband-noise maskers for three subjects (S1, circles with solid lines; S2, inverted triangles with dotted lines; and S3, squares with dashed lines). Masked thresholds (y-axis) are shown for 4.0 kHz (top panels), 1.5 kHz (middle panels), and 0.5 kHz (bottom panels) probe tones as a function of probe-tone location with respect to the masker onset (x-axis). Gray circles represent simultaneous masking, and filled symbols represent nonsimultaneous masking (backward or forward). The masker onset is designated as time zero. Open symbols on the ordinate between the ramped and damped panels represent each subject's probe-tone threshold in quiet.

an individual subject's previously measured thresholds. Average thresholds were calculated based on three runs unless the standard deviation exceeded 3 dB, in which case an additional run was obtained and included in the average. Thresholds for probe tones were obtained first in quiet and subsequently in the presence of ramped and damped maskers.

### D. Results and discussion

Figures 5–7 show temporal-masking patterns measured for the 10, 50, and 500 ms maskers, respectively. The greatest amount of masking occurred when the probe occurred within the masker, but the masker also produced backward and forward masking. Masking patterns for the 50 and 500 ms maskers (Figs. 6 and 7) followed the envelope of the masker. Simultaneous-masked thresholds for the ramped masker started at a low level and increased as the masker level increased, whereas thresholds for the damped masker

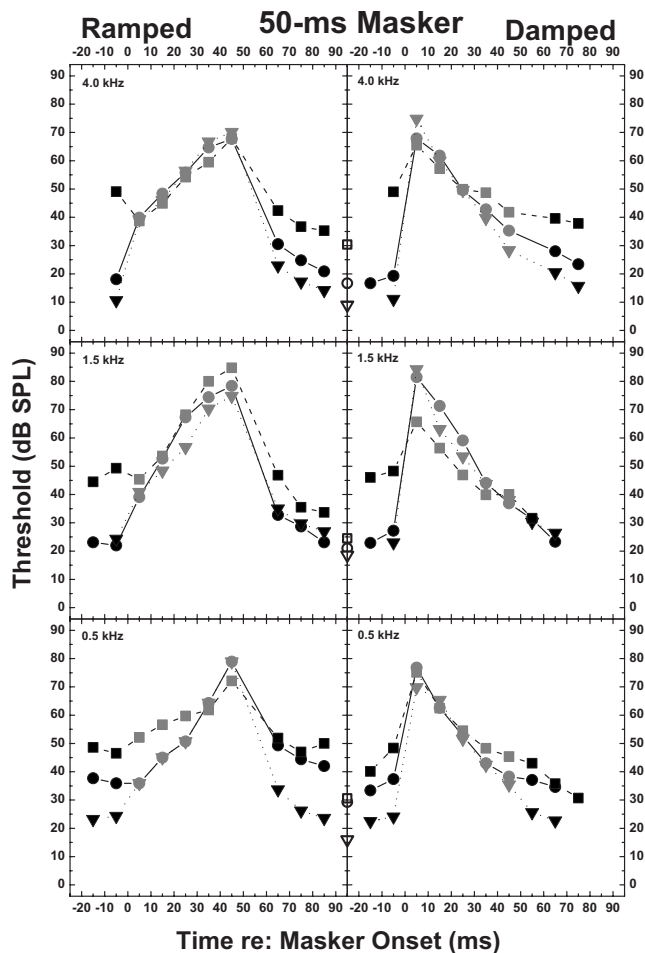


FIG. 6. Similar to Fig. 5, except that the masker duration was 50 ms.

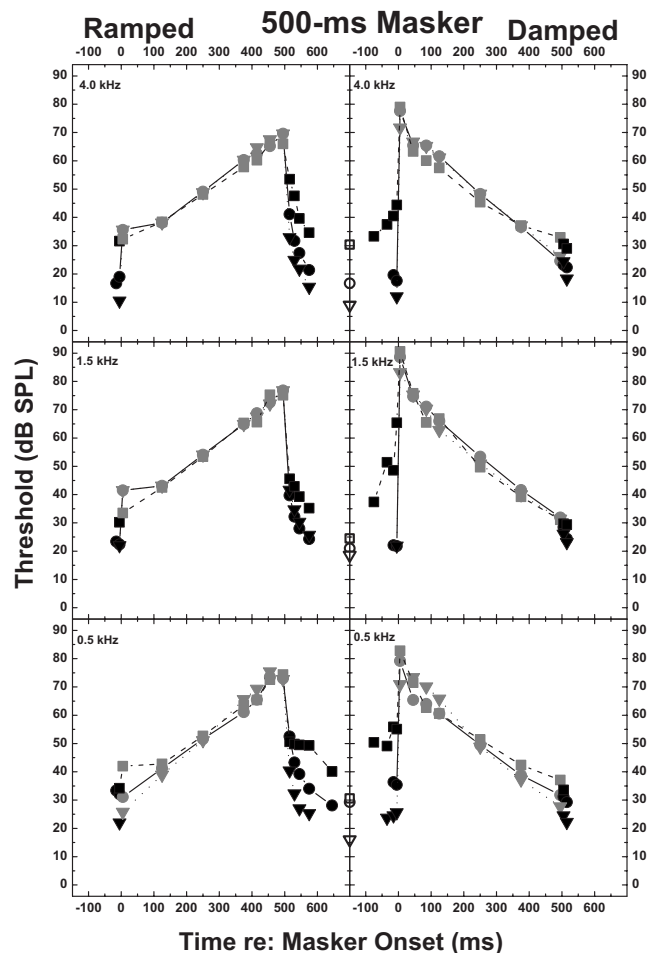


FIG. 7. Similar to Fig. 5, except that the masker duration was 500 ms.

started at a high level and decreased as the masker level decreased. This effect was not apparent for the 10 ms masker (Fig. 5) because the probe tone had the same duration as the masker. Nonetheless, the patterns for the 10 ms masker did show differences between forward and backward masking.

Forward- and backward-masked portions of the temporal-masking patterns are consistent with past research [e.g., see Fastl (1976) and Samoilova (1959)]. That is, backward-masked thresholds declined rapidly as the probe moved further from the onset of the masker, whereas forward-masked thresholds declined more gradually as the probe moved further from the end of the masker. Results for subject 3 (squares), however, showed some notable exceptions. This subject exhibited elevated backward-masked thresholds that declined very slowly compared to those of the other subjects, even upon repeated measurements. This was most apparent for the 50 ms masker (see Fig. 6); extensive training in the backward-masked conditions may have mitigated this disparity. Subject 3 reported receiving speech therapy as a child, but the nature of the problem and the treatment are unknown. Overall, the nonsimultaneous masked-thresholds followed the expected pattern, with the exception of some backward-masking results of subject 3.

Exponential functions were fitted to the nonsimultaneous masking data for each masker type (ramped or damped). The backward- and forward-masked functions

were expressed relative to the simultaneous-masked thresholds measured at the beginning (i.e., probe delay of 5 ms with respect to the masker onset for both ramped and damped maskers) and end of the masker (i.e., probe delays of 5 ms for the 10 ms masker, 45 ms for the 50 ms masker, and 495 ms for the 500 ms masker), respectively. An exponential-decay function was used to determine the time constant that produced the smallest sum-of-squared deviations about the backward- and forward-masking data for each subject [e.g., see Dolan and Small (1984), Hartley *et al.* (2000), Fastl (1976), Kidd and Feth, (1982), and Zwicker (1977, 1984)]. The fitted function was

$$Y = [(\theta_m - \theta_q)e^{-(t/\tau)}] + \theta_q, \quad (2)$$

where  $\theta_m$  is the masked threshold at a particular probe delay,  $\theta_q$  is the threshold in quiet,  $t$  is the probe delay with respect to the masker onset, and  $\tau$  is a time constant. The value of  $\tau$  producing the best fit for each masker type (ramped or damped) and for each masking condition (forward or backward) was used to fit a line to the data using Eq. (2). Again, these fits were expressed relative to the simultaneous masked thresholds at the beginning and end of the masker for the backward- and forward-masked functions, respectively. The best-fitting value of  $\tau$  was relatively short (<11 ms), although in some instances longer time constants were found.

TABLE III. Predicted (columns 2 and 3) and measured (column 4) perceived duration ratios of damped and ramped sounds with durations of 10, 50, and 500 ms (column 1). The ratios shown indicate what factor the physical duration of the damped sound needs to be greater than that of the ramped sound to achieve equal perceived duration (see text for details).

Stimulus duration (ms)	Predicted		Measured
	5 dB SL criterion	10 dB SL criterion	Matching task
10	1.10	1.19	1.25
50	1.19	1.22	1.64
500	1.08	1.06	1.20

There were four occasions out of 18 when the best-fitting line was horizontal, showing no exponential decay. All of the horizontal fits occurred for the backward-masked condition. Subject 1 accounted for two cases of 50 ms ramped masker, 0.5 kHz probe and 500 ms ramped masker, 0.5 kHz probe, and subject 3 accounted for two cases of 50 ms ramped masker, 4.0 kHz probe and 50 ms ramped masker, 1.5 kHz probe. Horizontal fits to the data of subject 1 were most likely due to data points for these conditions being within about 6 dB of absolute threshold and to inherent measurement noise. Horizontal fits for subject 3 were likely due to the extended backward masking described earlier. With these exceptions, exponential-decay functions provided an adequate description of the backward- and forward-masked data.

### 1. Predicting perceived duration

The exponential functions fitted to each subject's data were used to find the probe delays that produced backward- and forward-masked thresholds of 5 and 10 dB SL for each masker type and duration for each probe-tone frequency. The difference in time between these points represents the predicted subjective duration (Fastl, 1977; Fastl and Zwicker, 2007) under the assumption that nonsimultaneous masking is, at least in part, a reflection of the buildup and decay of excitation within the central auditory system. For example, using a 5 dB SL criterion for subject 1 with a 50 ms ramped masker and a 4.0 kHz probe tone results in backward- and forward-masked duration extensions of 3.1 and 27.9 ms, respectively. These values are then added to the physical duration of the masker (50 ms) to produce a total predicted perceived duration of about 81 ms. The medians of the predicted subjective durations across probe-tone frequency and subject for the ramped- and damped-broadband noises were used to determine the damped/ramped ratios listed in Table III. The values indicate that ramped sounds are predicted to sound longer than damped sounds for each stimulus duration, as the damped/ramped ratios are greater than 1. However, the difference predicted from the masking patterns is less than the obtained difference, which indicates either that there is an influence of central processes in the estimation of the duration of rising and falling intensity sounds or that the underlying assumptions regarding persistence of excitation and subjective duration are incorrect.

### 2. Loudness predictions

Masked thresholds obtained from the temporal-masking patterns can provide insights into the possible link between loudness and excitation for ramped and damped sounds as loudness can be predicted from these patterns (Fastl, 1984; Moore and Glasberg, 1996; Moore *et al.*, 1997; Zwicker and Scharf, 1965). Two recent implementations of loudness models for dynamically changing sounds represent extensions of an earlier model by Zwicker (1977). Both models base their predictions on neural excitation in response to a stimulus, and both assume that the loudness impression increases quickly with the onset of a sound and decays more slowly when a sound is terminated. Chalupper and Fastl (2002) modeled this aspect of loudness for dynamically changing sounds using forward masking to estimate the persistence of excitation, whereas Glasberg and Moore (2002) attributed it to the possible persistence of sensory information at some level in the auditory system that may be related to forward masking.

The loudness models for dynamically changing sounds of Glasberg and Moore (2002) and Chalupper and Fastl (2002) were used to predict loudness differences for the 10, 50, and 500 ms ramped and damped stimuli in our study. The models used the waveforms for our stimuli as input. The detailed descriptions of these models can be found elsewhere (Glasberg and Moore, 2002; Chalupper and Fastl, 2002). In Glasberg and Moore's (2002) model, the basic steps for the first stage include (1) filtering that represents the transfer functions for the outer and middle ear,<sup>1</sup> (2) estimation of the short-term spectrum at 1 ms intervals, (3) derivation of the pattern of excitation from the spectrum, (4) transformation of the excitation pattern to units of specific loudness, and (5) integration of the area under the specific loudness pattern. According to the authors, this first stage of their model yields instantaneous loudness, a quantity unavailable for conscious perception that possibly corresponds to neural activity. This calculation of instantaneous loudness is used to derive short-term loudness and long-term loudness, perceptual quantities that are based on the summation of these instantaneous values. The attack and release times used to calculate the short-term and long-term loudness values implement a temporal integration mechanism resembling an automatic gain control circuit. Short-term loudness estimates are derived directly from temporal integration of the instantaneous loudness values, whereas long-term loudness estimates are derived from temporal integration of short-term loudness values. Short-term and long-term loudness values both have quick attack times and slower release times, but the long-term loudness has more gradual attack and release times than does short-term loudness. Short-term loudness corresponds to loudness perceived at a particular instant in time, whereas long-term loudness represents the overall loudness impression. The predictions for loudness differences between ramped and damped stimuli from the present study using the models of Glasberg and Moore (2002) as well as that of Chalupper and Fastl (2002) are shown in Table IV, along with the results from the loudness-matching experiment. Glasberg and Moore's (2002) model predicted only a slight short-term loudness difference for the 10 ms stimuli and small differ-

TABLE IV. Predicted and measured level differences (dB) for equal loudness for ramped and damped sounds with durations of 10, 50, and 500 ms (column 1). Predicted values were derived from the ongoing estimates of short- and long-term loudness (columns 2 and 3) provided by the model of Glasberg and Moore (2002) at 1 ms intervals over and extending beyond the physical duration of the stimulus. The predictions are based on the difference between the peak values within each loudness type (short-term and long-term loudness) for a given duration. Loudness differences derived from the model of Chalupper and Fastl (2002) are provided in column 4. Positive values indicate that the ramped sound was predicted or measured (column 5) to be louder than the damped sound of equivalent duration.

Stimulus duration (ms)	Predicted			Measured Matching task
	Glasberg and Moore (2002)		Chalupper and Fastl (2002)	
	Short-term loudness	Long-term loudness	Estimated loudness	
10	+0.25	+0.05	+0.13	+1.04
50	+3.05	+1.15	+1.59	+1.97
500	+2.00	+2.40	+2.76	+2.33

ences in short-term loudness, consistent with the data, for the 50 and 500 ms stimuli. The long-term loudness values are also reasonably close to the matching results, especially for the 500 ms stimuli. The subjects may be using the short-term loudness values when dealing with the shorter-duration stimuli and the long-term loudness value when dealing with the long-duration stimuli (Glasberg and Moore, 2002). The predicted level differences required for equal loudness derived using the model of Chalupper and Fastl (2002) are also close to the average matching results.

These loudness models are based on the assumption that neural activity persists after a sound terminates. To determine if such an assumption is necessary, that is, to determine if excitation occurring during the time course of the maskers could explain the observed loudness differences, we inspected the temporal-masking patterns just during the on-time of the masker. The temporal-masking patterns for simultaneous segments averaged across listeners are shown in Fig. 8 for noise bands with ramped and damped envelopes and durations of 10, 50, and 500 ms. For an easy visual comparison at the same masker levels, the masking patterns for the ramped noise bands are reversed in time. The upper panels show group-mean data averaged across frequency for each of the durations. The middle panels and lower panels show group-mean data for signal frequencies of 4.0 kHz and 0.5 kHz, respectively. The ramped noise bands produced, on average, more masking than the damped noise bands. This trend was also evident in the data obtained using the 1.5 kHz probe tone (not shown). Given that loudness models based on masking patterns are derived under the assumption that the amount of masking is related to the level of excitation produced by a stimulus, these results are consistent with ramped noise bands being louder than damped noise bands. Predictions of the absolute difference in level that is required for equal loudness perception, however, are not accurate as a large effect of envelope would be expected for the 10 ms stimuli (see Fig. 8, top panel, left column), which was not observed in the matching data.

## V. GENERAL DISCUSSION AND CONCLUSIONS

This study adds to the growing body of literature indicating that ramped sounds appear louder and longer than

damped sounds with the same physical duration and intensity. In the present study, quantitative predictions of loudness and duration were made using models that are consistent with the notion that the internal representation of a sound persists after it is terminated. Modeling of duration used the extent of the temporal-masking pattern. Loudness values were predicted using models proposed by Glasberg and Moore (2002) and Chalupper and Fastl (2002). The loudness

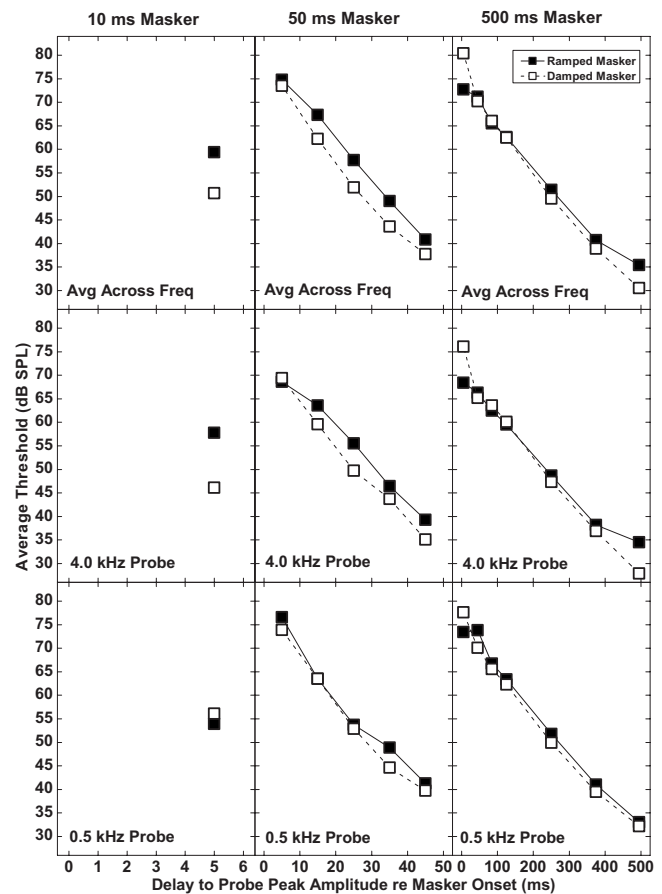


FIG. 8. Temporal-masking patterns for the 10 ms (left column), 50 ms (middle column), and 500 ms (right column) ramped (filled symbols) and damped (open symbols) broadband-noise maskers measured over only the physical duration of each stimulus. For ease of comparison, the patterns obtained for the ramped masker were reversed in time.

models based on persistence were consistent with the behavioral data, but the model for duration underpredicted the asymmetry in the behavioral data.

An alternative possibility to the persistence of excitation explanation is that sounds with ramped envelopes invoke more neural activity while the sound is on than sounds with damped envelopes. The amount of neural activity generated by a particular stimulus can be inferred from simultaneous masking. Thresholds for brief signals obtained during the time course of ramped noise bands were often higher than those for damped noise bands when the stimulus was at equal levels. Higher thresholds are consistent with a greater amount of excitation, a result which could be associated with the perceptual differences for ramped and damped sounds. This finding of higher thresholds during ramped sounds than damped sounds also suggests that forward masking of the tail of damped sounds by earlier occurring more intense portions of the stimulus is likely not responsible for the differences in loudness and duration that are perceived. It is not clear how these threshold differences, illustrated in Fig. 8, are converted into duration and loudness differences. The small ramped-damped threshold differences for a probe tone occurred only at the onset of the 500 ms noise band, yet a substantial difference was observed for the 10 ms noise band. This further complicates the analysis because the threshold asymmetry does not correspond to the magnitude of the observed loudness differences.

Threshold differences in the masking patterns for ramped and damped sounds during the time course of the masker may be related to overshoot. Overshoot is defined as poorer detection (i.e., increased threshold) of a brief tone at the onset of a noise masker than when the tone is delayed relative to the masker onset. Adaptation of neural fibers is a popular explanation for overshoot (Bacon and Takahashi, 1992), and the continuously rising intensity of a ramped sound could be processed by the auditory system akin to a set of continuous signal onsets, thereby limiting or reducing the effect of adaptation. The overall mean data for 10 and 50 ms stimuli (Fig. 8, top panels, left and middle columns) show a trend that would fit this supposition. The same outcome was not apparent in the data obtained for the 500 ms stimulus. The change in intensity over the 500 ms stimulus may be too gradual to cause this prolonged overshootlike effect and could be linked to the smaller decibel difference required to equate loudness for ramped and damped stimuli at longer durations. There is no established method for predicting loudness from these threshold differences, whereas the loudness models based on the persistence of excitation predict the loudness-matching results across durations successfully.

While models based on the persistence of excitation provide successful prediction differences in loudness between ramped and damped sounds, the same cannot be said for differences in subjective duration. The modeling results for subjective duration obtained using persistence of excitation, as estimated from nonsimultaneous masking data, underestimated the perceived difference in the duration of these stimuli (see Table III). Prior studies of subjective duration for ramped and damped sounds have reported similar findings

(DiGiovanni and Schlauch, 2007; Grassi and Darwin, 2006; Schlauch *et al.*, 2001a, 2001b). DiGiovanni and Schlauch (2007) found that the disparity in subjective duration judgments was largely due to subjects ignoring the portion of a damped signal that was perceived as echoic. The subjects in their study produced less disparate results for sounds with ramped and damped envelopes when specifically instructed to include all portions of the sound when making their judgments. This finding fits the decay suppression (i.e., perceptual constancy) hypothesis proposed by Stecker and Hafter (2000). Accordingly, listeners typically tend to ignore a portion of the decay of sounds with damped envelopes when judging duration. However, when listeners are instructed to include all aspects of ramped and damped sounds in their judgments of duration, the ramp-damp difference is smaller but a difference still remains (DiGiovanni and Schlauch, 2007). This difference is consistent with the predictions of persistence based on nonsimultaneous masking.

A significant context effect is observed in perceptual studies of sounds with ramped and damped envelopes, including the present study. The presence of a ramped sound prior to a damped sound produces a greater difference in subjective duration than when the order is reversed. However, the opposite is true for loudness judgments. An effect of stimulus context on loudness judgments was also reported and discussed by Stecker and Hafter (2000). They found, similar to the results reported here (see Fig. 4), that loudness judgments between ramped and damped noises were more disparate when a damped sound preceded a ramped sound than vice versa. Can perceptual constancy, as described by Stecker and Hafter (2000), help explain these seeming disparate results for judgments of duration and loudness based on biasing of responses due to stimulus order? Consider the experimental condition from the present study in which the stimulus with a ramped envelope was presented first. The ramped signal ended at a relatively high intensity, yet no possible interpretation of a reverberant (e.g., echoic) environment was allowed by the acoustic information available to that point. When the subsequent damped signal was presented for comparison, two likely acoustic-based environmental interpretations were plausible (1) the environment had not changed, the setting was nonreverberant, and a more “veridical” (i.e., with little effect of decay suppression) version of the second signal should be used to compare the two stimuli, or (2) the listener had entered a new reverberant setting between the sounding of the signals and the second stimulus should be processed in the “constancy” mode in order to account for the influence of “echoic” information upon perception. Estimates of loudness and duration for the ramped and damped sounds using environmental interpretation 1 would be less different from each other as the effect of decay suppression upon the damped stimulus would be limited. This situation fits the trends evident in the loudness-matching data but not in the duration-matching data. Using environmental interpretation 2, the loudness and duration estimates would be more disparate, from those obtained using interpretation 1, as a portion of the persistence of excitation of the damped stimulus would be ignored. This second interpretation follows the trend seen in the duration-matching

data but not that for loudness matching. Thus, the explanation for this context effect remains a mystery.

In summary, the temporal-masking patterns measured in this study do not provide unequivocal support for the persistence, adaptation, or perceptual constancy approaches for modeling perceptual differences between ramped and damped sounds. That stated, the loudness models based on the persistence approach were able to yield quantitative predictions that were consistent with the average differences in the loudness data for each of the durations used. The success of the persistence approach for modeling loudness has been reported for a variety of stimuli (Chalupper and Fastl, 2002; Glasberg and Moore, 2002). The failure of the persistence approach to account for subjective duration as well as the divergent effects of context on loudness and duration suggest that different underlying mechanisms control these two facets of auditory perception.

## ACKNOWLEDGMENTS

This work was supported in part by NIH/NIDCD Grant No. R29 DC01542. We thank Edward Carney for his assistance in stimulus generation and Brian Glasberg and Brian Moore as well as Hugo Fastl for running our ramped and damped stimuli through their models. We also thank the associate editor and the two reviewers for their helpful comments.

<sup>1</sup>A diffuse field filter was used to model the response for our stimuli, and although our behavioral data were collected using insert earphones, any differences between the output of the models and the behavioral data are expected to be trivial. Alternate filters that simulate the transfer functions for a variety of listening environments can be implemented in the loudness models of Glasberg and Moore (2002) and Chalupper and Fastl (2002).

- Bacon, S. P., and Takahashi, G. A. (1992). "Overshoot in normal-hearing and hearing-impaired subjects," *J. Acoust. Soc. Am.* **91**, 2865–2871.
- Chalupper, J., and Fastl, H. (2002). "Dynamic loudness model (DLM) for normal and hearing-impaired listeners," *Acust. Acta Acust.* **88**, 378–386.
- DiGiovanni, J. J., and Schlauch, R. S. (2007). "Mechanisms responsible for differences in perceived duration for rising-intensity and falling-intensity sounds," *Ecological Psychol.* **19**, 239–264.
- Dolan, T. G., and Small, A. M. (1984). "Frequency effects in backward masking," *J. Acoust. Soc. Am.* **75**, 932–936.
- Fastl, H. (1976). "Temporal masking effects: I. Broadband noise masker," *Acustica* **35**, 287–302.
- Fastl, H. (1976/1977). "Temporal masking effects: II. Critical band noise masker," *Acustica* **36**, 317–330.
- Fastl, H. (1977). "Subjective duration and temporal masking patterns of broadband noise impulses," *J. Acoust. Soc. Am.* **61**, 162–168.
- Fastl, H. (1979). "Temporal masking effects: III. Pure tone masker," *Acustica* **43**, 282–294.
- Fastl, H. (1984). "Folgedrosselung von sinustönen durch breitbandrauschen: Messergebnisse und modellvorstellungen (Temporal partial masking of pure tones by broad-band noise)," *Acustica* **54**, 145–153.
- Fastl, H., Büeler, R., and Fruhmann, M. (2002). "Different implementations of a model for subjective duration," *Fortschritte der Akustik, DAGA 2002* (Dt. Gessel. für Akustik e. V., Oldenburg), pp. 470 and 471.
- Fastl, H., and Zwicker, E. (2007). "Subjective duration," *Psychoacoustics: Facts and Models*, 3rd ed. (Springer-Verlag, Berlin), pp. 265–269.
- Gabriel, B., Kollmeier, B., and Mellert, V. (1997). "Influence of individual listener, measurement room, and choice of test-tone levels on the shape of equal-loudness level contours," *Acust. Acta Acust.* **83**, 670–683.
- Glasberg, B. R., and Moore, B. C. J. (2002). "A model of loudness applicable to time-varying sounds," *J. Audio Eng. Soc.* **50**, 331–341.
- Grassi, M., and Darwin, C. J. (2006). "The subjective duration of ramped and damped sounds," *Percept. Psychophys.* **68**, 1382–1392.
- Hartley, D. E. H., Wright, B. A., Hogan, S. C., and Moore, D. R. (2000). "Age-related improvements in auditory backward and simultaneous masking in 6- to 10-year-old children," *J. Speech Lang. Hear. Res.* **43**, 1402–1415.
- Irino, T., and Patterson, R. D. (1996). "Temporal asymmetry in the auditory system," *J. Acoust. Soc. Am.* **99**, 2316–2331.
- Kidd, G., and Feth, L. L. (1982). "Effects of masker duration in pure-tone forward masking," *J. Acoust. Soc. Am.* **72**, 1384–1386.
- Levitt, H. (1971). "Transformed up-down methods in psychoacoustics," *J. Acoust. Soc. Am.* **49**, 467–477.
- Maier, J. X., Neuhoff, J. G., Logothetis, N. K., and Ghazanfar, A. A. (2004). "Multisensory integration of looming signals by Rhesus monkeys," *Neuron* **43**, 177–181.
- Miller, G. A. (1948). "The perception of short bursts of noise," *J. Acoust. Soc. Am.* **20**, 160–170.
- Moore, B. C. J., and Glasberg, B. R. (1996). "A revision of Zwicker's loudness model," *Acust. Acta Acust.* **82**, 335–345.
- Moore, B. C. J., Glasberg, B. R., and Baer, T. (1997). "A model for the prediction of thresholds, loudness, and partial loudness," *J. Audio Eng. Soc.* **45**, 224–240.
- Neuhoff, J. G. (1998). "Perceptual bias for rising tones," *Nature (London)* **395**, 123–124.
- Neuhoff, J. G. (2001). "An adaptive bias in the perception of looming auditory motion," *Nature (London)* **398**, 673–674.
- Nuttall, A. H. (1981). "Some windows with very good sidelobe behavior," *IEEE Trans. Acoust., Speech, Signal Process.* **29**, 84–91.
- Patterson, R. D. (1994a). "The sound of a sinusoid: Spectral models," *J. Acoust. Soc. Am.* **96**, 1409–1418.
- Patterson, R. D. (1994b). "The sound of a sinusoid: Time-interval models," *J. Acoust. Soc. Am.* **96**, 1419–1428.
- Plomp, R. (1964). "Rate of decay of auditory sensation," *J. Acoust. Soc. Am.* **36**, 277–282.
- Ryan, T. A. (1959). "Comments on orthogonal components," *Psychol. Bull.* **56**, 394–396.
- Samoilova, I. K. (1959). "Masking of short tone signals as a function of the time interval between masked and masking sounds," *Biophysika* **4**, 550–558 [Biophysics (Engl. Transl.) **4**, 44–52].
- Schlauch, R. S., DiGiovanni, J. J., Olson, K. L., and Donlin, E. E. (2001a). "Do listeners include the echo portion of a damped sound when judging its duration?," *J. Acoust. Soc. Am.* **109**, 2465.
- Schlauch, R. S., Ries, D. T., and DiGiovanni, J. J. (2001b). "Duration discrimination and subjective duration for ramped and damped stimuli," *J. Acoust. Soc. Am.* **109**, 2880–2887.
- Schlauch, R. S., Ries, D. T., DiGiovanni, J. J., Elliot, S., and Campbell, S. (1998). "Intensity discrimination of ramped and damped tones," *Proceedings of the 16th International Congress on Acoustics and 135th Meeting of Acoustical Society of America Vol. 2*, pp. 885–886.
- Stecker, G. C., and Hafter, E. R. (2000). "An effect of temporal asymmetry on loudness," *J. Acoust. Soc. Am.* **107**, 3358–3368.
- Zwicker, E. (1977). "Procedure for calculating loudness of temporally variable sounds," *J. Acoust. Soc. Am.* **62**, 675–682.
- Zwicker, E. (1984). "Dependence of post-masking on masker duration and its relation to temporal effects in loudness," *J. Acoust. Soc. Am.* **75**, 219–223.
- Zwicker, E., and Scharf, B. (1965). "A model of loudness summation," *Psychol. Rev.* **72**, 3–26.
- Zwislocki, J., Pirodda, E., and Rubin, H. (1959). "On some poststimulatory effects at the threshold of audibility," *J. Acoust. Soc. Am.* **31**, 9–14.