Kent State University Libraries (KSU)
Interlibrary Loan

ILLiad TN: 199501 ARIEL/ODYSSEY

Borrower: OUN

In Process Date:20071017

Lending String: MIA,*KSU,BUF,BUF,VWM

Patron: DiGiovanni, Jeffrey, DEPT;Hearing & Speech Scienc

Journal Title: Ecological psychology ; a publication of the International Society for Ecological Psychology.

Volume: 19 **Issue:**
Month/Year: 2007**Pages:** 239-264

Article Author:

Article Title: DiGiovanni JJ & Schlauch RS; Mechanisms responsible for differences in perceived duration for rising-intensity and falling-intensity sounds

Imprint: Hillsdale, N.J. ; Lawrence Erlbaum Assoc

ILL Number: 36095066

Call #:

Location: C

Charge
Maxcost: $0.00IFM

Shipping Address:
Ohio University
Library Bldg. 5-CG, ILL
30 Park Place
Athens OH 45701-2978

Fax: 740-593-2959
Ariel: ARIEL.LIBRARY.OHIOU.EDU

KENT STATE
UNIVERSITY

ECOLOGICAL PSYCHOLOGY, *19*(3), 239–264

Mechanisms Responsible for Differences in Perceived Duration for Rising-Intensity and Falling-Intensity Sounds

Jeffrey J. DiGiovanni
Ohio University

Robert S. Schlauch
University of Minnesota

Sounds that are equivalent in all aspects except for their temporal envelope are perceived differently. Sounds with rising temporal envelopes are perceived as louder, longer, and show a greater change in loudness throughout their duration than sounds with falling temporal envelopes. Stecker and Hafter (2000) proposed that participants ignore the decay portion of sounds with falling temporal envelopes to account for observed loudness differences, but there is no empirical evidence support this hypothesis. To test this idea, two duration-matching experiments were performed. One experiment used broadband noise and the other natural stimuli. Different groups of participants were given different instruction sets asking them to (1) simply match the duration or (2) include all aspects of the sounds. Both experiments produced the same result. The first instruction set, which represented participants' natural biases, yielded shorter subjective durations for sounds with falling temporal envelopes than for sounds with rising temporal envelopes. By contrast, asking participants to include all aspects of the sounds significantly reduced the size of the asymmetry in subjective duration, a result that supports Stecker and Hafter's hypothesis. This segregation of the stimulus at the perceptual level is consistent with observed asymmetries in loudness change and overall loudness for sounds with rising and falling temporal envelopes, but it does not account for the entire effect. The remaining portion of

---

Correspondence should be addressed to Jeffrey J. DiGiovanni, W222 Grover Center, School of Hearing, Speech and Language Sciences, Ohio University, Athens, OH 45701. E-mail: digiovan@ohio.edu

the effect, after considering biases due to instructions, is not likely a result of adaptation but could be associated with persistence. The amount of persistence was inferred from behavioral masking data obtained for these sounds.

Numerous studies have demonstrated differences in duration and loudness for sounds with rising and falling intensity (Maier, Neuhoff, Logothetis, & Ghazanfar, 2004; Neuhoff, 1998, 2001; Patterson, 1994a, 1994b; Patterson & Irino, 1998; Schlauch, Ries, & DiGiovanni, 2001; Seifritz et al., 2002). The primary finding of these studies is that sounds with rising temporal envelopes are perceived as louder and having a greater change in loudness than sounds with falling temporal envelopes (Neuhoff, 1998, 2001).

Neuhoff (2001) hypothesized that a bias, or a greater judgment of intensity change, toward sounds that rise in intensity provides an organism with a selective advantage. A rising temporal envelope represents an acoustic correlate for an approaching sound whereas a falling temporal envelope corresponds to a sound moving away (Neuhoff, 1998). According to Neuhoff's hypothesis, the selective advantage for rising-intensity sounds is gained by enabling the listener to better prepare for an approaching object. Furthermore, Ghazanfar, Neuhoff, and Logothetis (2002) have shown this same preferential bias toward risings sounds to be common for nonhumans as well.

Natural sounds are often short in duration (less than 1 sec) and produced with a gradually falling temporal envelope. This class of sounds is characteristic of the envelopes for many impulsive sounds, ones produced when something is struck. This category contains not only environmental sounds (a breaking twig or a falling branch) but also many animal vocalizations, including many speech sounds (e.g., stop consonants).

By contrast, sounds with rising temporal envelopes are associated with an approaching sound source as noted by Neuhoff (1998). Sounds with gradually rising and gradually falling temporal envelopes are also common as alarm vocalizations for many small mammals and birds (Bradbury & Vehrencamp, 1998). Selective evolutionary pressures may have led to alarm vocalizations with gradual onsets and offsets in these small animals. In particular, an alarm signal that is designed to convey the message for others to flee a predatorial situation needs to be difficult to localize in order to protect the location of the signaler. The combination of slow onsets, slow offsets, and pure-tone signals all make localization difficult. Therefore, these types of stimuli protect the location of these small animals and birds while allowing them to alarm others (Bradbury & Vehrencamp, 1998). This explanation might also lead to sounds with rising temporal envelopes as being more salient than more commonly occurring sounds with abruptly rising and gradually falling temporal envelopes. Although animal alarms typically begin and end with a gradual change in amplitude, they begin with the less common rising-intensity envelope, which would likely be perceived as an alerting stimulus. Neuhoff (2001) has well described how rising stimuli appear as if there were greater stimulation to the auditory system than equivalent falling stimuli. It has also been shown that rising and falling sounds indicate motion and that rising sounds indicated motion of greater distance than equivalent falling sounds (Seifritz et al., 2002). It has been argued that, taken together, this neural processing that leads to the greater stimulation of rising sounds helps provide warning for looming sounds (Neuhoff, 2001; Seifritz et al., 2002). This warning, in turn, provides the adaptive advantage to respond to approaching, or looming, sounds. The high priority of rising intensity sounds has been found at the cortical level, although it is not clear if the cortical response is reflective of a simple sensory process, coding of the acoustic signal carried through the auditory pathway, or a perceptual process requiring specialized computational and/or integrative processing (Hall & Moore, 2003; Seifritz et al., 2002).

Several prior studies have documented the salience of rising-intensity sounds. Pairs of stimuli used to study differences in subjective duration and loudness for sounds with rising and falling temporal envelopes are produced in a laboratory by simply time-reversing a sound with a falling temporal envelope. This simple manipulation produces a sound with a rising envelope with the identical duration and long-term amplitude spectrum as the original sound with a falling temporal envelope. Studies employing these stimuli have found differences in loudness change (Neuhoff, 1998, 2001; Seifritz et al., 2002), overall loudness (Stecker & Hafter, 2000) and perceived duration (Schlauch, Ries, DiGiovanni, Elliot, & Campbell, 1998; Schlauch et al., 2001; Grassi & Darwin, 2001). These studies found that sounds with rising temporal envelopes are judged to change more in loudness, to be louder, and to have a longer duration than their time-reversed control stimulus with a falling temporal envelope.

Listeners have reported that pure tones with rising intensity have a more "sinusoidal" percept whereas similar sounds with a falling intensity sound more "hollow" and "percussive" (Patterson, 1994a, 1994b). This distinction is considered a difference in timbre. Many authors define timbre as the attribute that enables a listener to distinguish between complex sounds that have the same loudness, pitch, and duration (ANSI S3.2, 1995; Grey, 1977; Plomp, 1970). For instance, saxophone and a trumpet playing the same note are easily recognized as different instruments and it is their timbre that enables this distinction when they are matched for pitch, loudness, and duration. This definition, based on exclusion (Risset & Wessel, 1999), assumes that sounds having different timbres are matched for duration, the topic of this article, but prior studies have not typically specified whether the duration should be the subjective duration or the physical duration of the sounds. Pitch and loudness are subjective quantities and one might assume that the subjective duration should be used rather than physical duration to remain consistent. One reason a distinction has not been made previously is perhaps that investigators did not realize that perceived duration may differ from the physical duration. Given the role duration that plays in timbre judgments and the finding that sounds with rising and falling temporal envelopes are judged to have different timbres, it is important to understand the mechanism underlying differences in perceived duration for this class of sounds.

Stecker and Hafter (2000) speculated about the mechanism underlying differences in the loudness of sounds with rising and falling intensity envelopes. They proposed a perceptual explanation to account for the loudness difference, and this explanation is also consistent with differences in subjective duration for these sounds. According to their theory, listeners ignore a segment of the decay portion, the portion representing a decreasing sound pressure level, of falling-intensity sounds because they consider it an echo or reverberation, intrinsic or extrinsic acoustic reflections increasing the overall duration of the sound. If listeners ignore the decay portion of a sound with a falling temporal envelope it will be perceived to be shorter and softer than a sound with a rising temporal envelope of the same physical duration.

Stecker and Hafter's (2000) explanation for differences in subjective duration and loudness between sounds with rising and falling temporal envelopes is based on the idea that listeners can segment the attack (the portion of the sound during its onset) and release, or decay, portions of a sound to glean information about the identity of a sound and its physical characteristics. In their paper, they state that for many sounds the driving force provides information about the attack portion of a sound (e.g., striking a drum or a cymbal) whereas filtering characteristics determine the decay. The filtering characteristics could include several coupled filters. In the case of speech produced in a highly reverberant environment, these filters would include the vocal tract of a talker and the characteristics of the room where the utterance occurred.

For judgments regarding the duration of a sound with an abrupt attack and a gradual release (a falling temporal envelope), listeners can segment the decay portions separately from earlier portions in some extreme conditions. Intuitively, it seems that the decay portion of a word spoken in a highly reverberant environment would not be included as part of the talker's original utterance. It follows that the judged duration may be different depending on whether the decay is considered as part of the judgment. Studies have also shown that listeners are able to judge room characteristics based on their reverberation times; this demonstrates that listeners are able to segment a sound in a manner to recover information about reverberant space (Barron & Marshall, 1981; Benade, 1976). For instance, longer reverberation times are associated with larger rooms. Another aspect of reverberant sound, the ratio of reverberant to diffuse energy, provides a critical cue for judging loudness of a sound source in enclosed spaces (Zahorik & Wightman, 2001).

The major goal of this study is to assess whether the explanation for the salience of rising-intensity sounds is a result of perceptual process whereby listeners actively ignore the decay portion of a sound or a sensory process whereby listeners judge the encoded stimulus as a single event. To address this question, we began by testing Stecker and Hafter's (2000) hypothesis directly by instructing participants to include different aspects of sounds when judging their duration. In the first experiment, participants matched the duration of broadband noises with rising and falling temporal envelopes to the same broadband noise with rectangular temporal envelopes (turned on and off abruptly). These matches were obtained for different instruction sets. One instruction set asked participants to simply match the durations of the sounds. A second instruction set asked participants to include all aspects of the sounds in their judgments. The data obtained using these highly controlled laboratory stimuli provided a baseline for the second experiment that used natural stimuli. Experiment 2 employed a drum strike and a single word spoken in a simulated, highly reverberant environment to determine if listeners are capable of separating the attack and decay portions when judging duration. Because an "echo" was added to the spoken word, a third instruction set was included in this experiment that asked participants to ignore an echo if one existed. A third experiment assessed whether persistence, which can be thought of as a continuation of a sound's neural representation after the sound has terminated, contributes to the observed salience of rising-intensity sounds. Persistence was inferred from behavioral thresholds obtained from non-simultaneous masking data. Rising-intensity and falling-intensity sounds were used as maskers and thresholds for a signal were measured as a function of time after the termination of the masker (and before it began).

EXPERIMENT 1: THE EFFECT OF INSTRUCTIONS ON DURATION-MATCHING JUDGMENTS OF NOISE BANDS WITH DIFFERENT TEMPORAL ENVELOPES

This experiment used broadband noise as the stimulus to examine differences instructions have on the subjective durations of sounds with rising and falling temporal envelopes. Broadband noise was used to provide stimulus control. The average spectrum of broadband noise is more uniform than that of natural stimuli, such as speech. Although the asymmetry in subjective duration and loudness is usually smaller for a broadband noise than it is for a tone or natural stimulus (e.g., Neuhoff, 1998, 2001; Schlauch et al., 2001), one of these previous studies reported a significant asymmetry in subjective duration for broadband noise with the rising and falling intensity (Schlauch et al., 2001). In this prior study, the stimulus with a rising intensity was judged to be longer than the stimulus with a falling intensity. The present experiment is intended to replicate this finding and to determine if the size of the asymmetry in subjective duration changes when listeners are given instructions to include all aspects of the sounds.

Method

Participants. Twenty young adults with normal hearing participated. All participants had auditory thresholds of 15 dB HL or better for octave intervals between 0.25 and 8.0 kHz. Informed consent was obtained from all participants before data collection began.

Stimuli. Broadband noise was generated using a custom-designed 16-bit D/A converter. The sampling rate was 26 kHz. Anti-aliasing filters were set to 10 kHz. Participants listened monaurally through headphones (Telephonics, Model TDH-49P). Sounds with falling temporal envelopes were generated using a simple exponential decay with a time constant set to one fifth the signal duration according to the following formula:

$$x(t) = e^{\frac{-5}{T}} \times w(t)$$

where t is the time in s, w is a rectangular-gated noise, and T is the duration of w. Sounds with rising temporal envelopes were created by simply reversing the order of the samples in the array for the sounds with falling temporal envelopes prior to output. Stimuli were presented at 80 dB SPL (peak) and each falling-intensity stimulus decayed 42 dB over its duration. The duration of these rising-intensity and falling-intensity sounds was 50 ms or 500 ms. Rectangular-gated (turned off and on abruptly) broadband noise served as the comparison stimulus. The rectangular-gated noise had a peak level of 80 dB SPL, and its duration was adjustable over a large range.

Procedure. Participants performed a duration-matching task. The standard and comparison stimuli were presented in succession separated by a 500-ms silence. After a 1.5-sec delay, the two stimuli were repeated. During each subsequent presentation, the duration of the comparison stimulus was changed according to the slider position. The participants used a slider on a box to change the duration of the comparison stimulus. The participants' task was to match the duration of the two sounds. Participants were instructed to use the entire range of adjustments by moving the slider from the low end (left side) to the high end (right side) before making any match. The participants were also instructed to bracket the adjustment above and below the point of equal perceived duration to reduce bias. Finally, when they reached a point of equal duration, participants were instructed to press a button on the box to end the block. Once this block of trials ended, the experimenter started a new one. Complete instructions can be found in the Appendix. The standard stimulus was fixed at either 50 or 500 ms.

A slide on a response box controlled the comparison stimulus and after each match the range of adjustments was varied to ensure that participants were not focusing in on a single point on the slide (Schlauch et al., 2001). For the 500 ms stimuli, the rising-intensity stimuli included a range of adjustment from a randomly selected shortest duration of 2 – 203 ms to a longest duration of 1480 ms for the ramped condition. For the falling-intensity condition, the range was 40 ms to a randomly selected 870 – 1480 ms. For the 50 ms stimuli, the comparison stimulus had a minimum duration of 2 ms and a maximum duration of a randomly selected duration between 150 and 250 ms. Prior to data collection, participants completed a rectangular versus rectangular condition. The standard was 50 ms for 10 matches and 500 ms for 10 matches. The order of the standard and comparison was counterbalanced.

Twenty participants were randomly assigned to one of two groups (10 participants per group). The first group received instructions to simply match the duration of the standard and comparison stimuli. The second group was told to include all aspects of the sounds when making their duration matches. These were the "null" and "all" instruction sets of Experiment 1. The complete instructions can be found in the Appendix. Participants were presented with the following conditions, all of which were broadband-noise based stimuli: (1) rising-intensity versus rectangular-gated and (2) falling-intensity versus rectangular-gated. For these conditions, the rectangular-gated noise was the comparison stimulus. For 5 matches the standard was presented first; for the other 5 matches the comparison was presented first.

Results

Figure 1 illustrates the results obtained for the duration-matching tasks. These box plots (Tukey, 1977) represent the difference in participants' adjustments from the standard duration for each condition. For the two groups of participants, the matched durations of rectangular-gated broadband noise to itself were within a few percent of the standard duration for both instruction sets. This demonstrates that participants could perform the task with precision. For the "null" instruction set, rising-intensity sounds and falling-intensity sounds yielded results similar to those from Schlauch et al. (2001), who used the same instruction set and durations of 50 ms and 200 ms. That is, both studies found that rising-intensity broadband noise was judged to be about equal in duration to the physical duration of the stimulus, but falling-intensity broadband noise was judged shorter, on average, than the physical stimulus duration. The ratio of the matched durations for noise bands with rising and falling temporal envelopes (rising-falling duration ratio) provides a useful summary of the data separate from the standard. Ratios are also appropriate for analysis because the just-noticeable difference for duration is proportional to the duration of the standard stimulus (Abel, 1972). Group-median rising-falling duration ratios examined in this manner are 1.46 and 1.85 for the 50 and 500 ms conditions, respectively. This is similar to the findings of Schlauch et al. (2001), whose participants yielded rising-falling duration ratios with average values of 1.45 and 1.63 for 50 and 200 ms conditions, respectively. This finding shows that noise bands with rising temporal envelopes are judged to be roughly 1-1/2 times longer or more than sounds with falling temporal envelopes.

Figure 1 shows that the median matches for rising-intensity broadband noise were nearly identical for instruction sets "null" and "all." By contrast, falling-

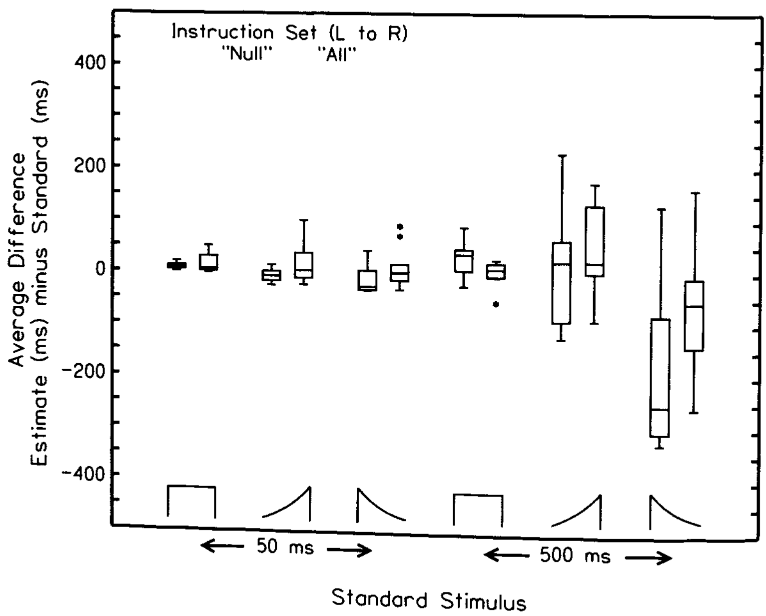

FIGURE 1 Box plot representations of the matching data for Experiment 2 for 10 participants. The data are represented as deviations of adjustments (ms) from the standard durations, which were 50 and 500 ms. The standard stimulus was a rising-intensity, falling-intensity, or rectangular-gated broadband noise. The adjustable, comparison stimulus was a rectangular-gated broadband noise. Above the labels on the abscissa for each stimulus are data for the "null" and "all" instruction sets, in order from left to right. The height of the boxes represents the inter-quartile range (IQR, the 25th percentile to the 75th percentile) and the horizontal lines within boxes display median values. Whiskers extend to the last data point or to the inner fence, which terminates at 1.5 times the IQR from the end of the box. Data points falling between the inner and outer fences (between 1.5 and 3.0 times the IQR) are represented by *. Data points falling beyond the

intensity broadband noise was judged as longer when participants were asked to consider the entire duration of the stimulus in their judgments than when their judgments were influenced by their natural bias in the "null" instruction set. The range of values that participants used for their matches for the two stimulus durations was much greater for the 500-ms conditions than for the 50-ms conditions, but Figure 2 illustrates that when rising-falling duration ratios are considered, an analysis that normalizes the data for these disparate durations, participants followed the same trend for both stimulus durations. Rising-intensity and falling-intensity broadband noise produce greater subjective differences for the "null" instruction set than for the "all" instruction set. This indicates that the underestimation of noise bands with falling temporal envelopes is minimized by asking participants to include all aspects of these sounds in their duration judgments. A

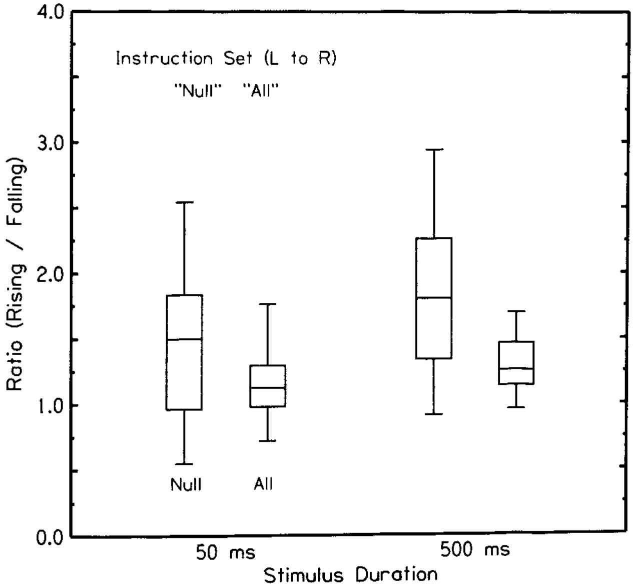

FIGURE 2 Box plot representations of the matching data for Experiment 2 for 10 participants. The data are plotted as rising-falling duration ratios for 50 and 500 ms noise bands for 2 instruction sets. The data are represented in the same manner as in Figure 1.

repeated measures analysis of variance (ANOVA) was completed using stimulus and instruction set as the within-subject factors for these rising-falling duration ratios. This analysis showed a significant effect of instruction set, $F(1, 18) = 5.90$; $p = .03$. A significant effect of duration was also found, $F(1, 18) = 4.8$; $p = .04$. There was not a significant instruction set by duration interaction. This analysis indicates that the rising-falling duration ratios were significantly smaller for the "all" instruction set versus the "null" instruction set. Furthermore, the same ratios for the 50 ms duration were significantly smaller for the 50 ms stimuli than for the 500 ms stimuli. However, the effect of different durations and instruction sets were independent from one another.

Discussion

Instructions had a significant effect on the judged duration of the noise bands with a falling temporal envelope. Prior studies comparing subjective durations for laboratory sounds with rising and falling temporal envelopes used our "null" instruction set (Schlauch et al., 2001; Grassi & Darwin, 2001). These instructions allow participants to approach the task with their inherent biases, which produce a large underestimation of the duration for a sound with a falling temporal envelope. When the instructions indicate that participants should include the entire sound in their judgments, differences in subjective durations for broadband noise with rising and falling temporal envelopes become smaller. The second experiment in this study replicates this finding for natural sounds.

EXPERIMENT 2: THE EFFECT OF INSTRUCTIONS ON DURATION JUDGMENTS OF NATURALLY PRODUCED RISING-INTENSITY SOUNDS AND FALLING-INTENSITY SOUNDS

Participants were required to match the subjective duration of a broadband noise to the subjective duration of four stimuli. The stimuli were a drum strike and single word, both played forward (giving a falling-intensity sound) and backward (giving a rising-intensity sound). In a control condition the duration of the broadband noise was matched to itself. The single word was recorded in near-anechoic conditions but was processed in a manner to simulate a highly reverberant environment. The broadband noise was just used as a comparison stimulus in this experiment.

Method

Participants. A total of 30 young adults (mean age: 21, range: 19–22 years) participated in this experiment. The participants were divided into three groups of 10. Each group participated in one of the three experimental conditions. All participants had auditory thresholds of 15 dB HL or better for octave intervals between 0.25 and 8.0 kHz. Informed consent was obtained from all participants before data collection began.

Stimuli. Three stimuli were used in this experiment: the word "do," a drum strike, and rectangular-gated broadband noise. All stimuli were presented using a custom-designed, 16-bit digital-to-analog converter. Participants listened monaurally through headphones (Telephonics, Model TDH-49P).

The word "do" was taken from a compact-disc recording from one of the Central Institute for the Deaf (CID) Everyday Sentences (Davis & Silverman, 1978) that was recorded in the Speech Perception Laboratory at the University of Minnesota. The sampling rate was 11,025 Hz. The drum strike had little perceptible nonsource generated reverberation (Santana, 1999). Therefore, the word "do" was digitally processed to simulate the reverberation of a large room using a digital signal processing algorithm available in a commercial sound editing software package (Syntrillium Software Corp., Cool Edit Pro). The modeled room had the dimensions of 26 x 38.2 x 16.8 m (width, depth, height) for a total volume of 16,686 m^3 and a width/depth ratio of 0.68. The total length was 3.51 sec with an attack time of 133 ms. The reverberation time, the time passed before the sound decayed 60 dB, was 1.75 sec. The duration of the word was truncated to 627 ms after the reverberation was added. The stimulus, which was 78 dB SPL at the onset, decayed by 42 dB at the point where it was truncated. An anti-aliasing filter was set to 5,000 Hz for the word "do" and 10,000 Hz for the drum strike. The waveforms are shown in Figure 3.

The word and the drum strike were terminated after a decay of 42 and 36 dB, respectively, for two reasons. First, the stimuli in Schlauch et al. (2001) were terminated after 42 dB decay, and we planned to compare our results to that study. Second, 36 to 42 dB of decay represents a significant reduction in the level of the sound. The level of the sounds just before they were truncated was only 20 dB SL. By choosing these levels and the amount of decay, we ensured audibility of the entire sound while maintaining a natural percept. Maintaining a suprathreshold level

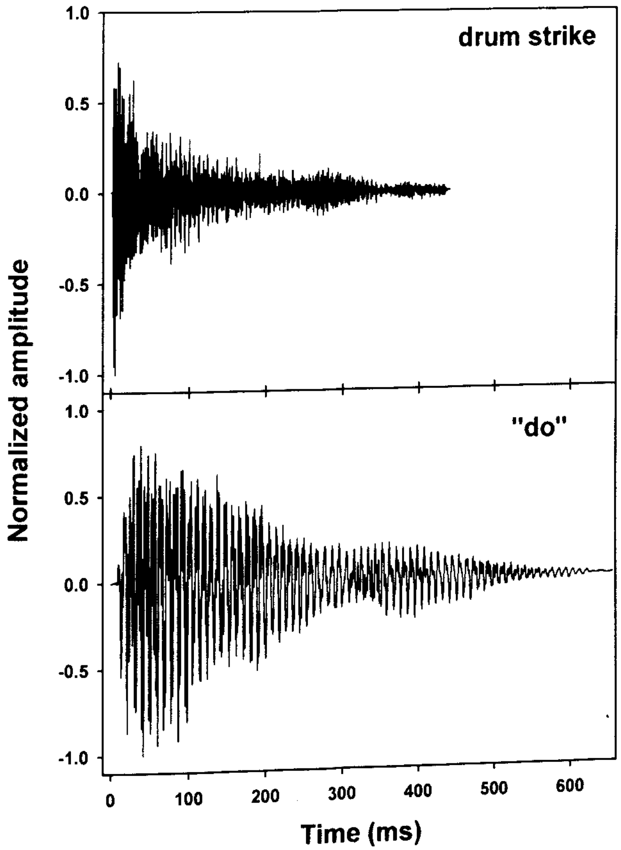

FIGURE 3 Waveforms for the drum strike and the word "do" are shown in the upper and lower panels, respectively.

for the entire duration is critical for evaluating possible mechanisms that underlie the asymmetries in subjective aspects of these sounds. This issue will be raised again in the general discussion.

The broadband noise was generated with a different sampling rate for each stimulus it was paired with: 11,025 Hz with the word, 22,050 Hz with the drum strike in order to match the sampling rates of the natural stimuli. The sampling rate of the broadband noise was 26,000 Hz when it was paired with itself. The noise was low-pass filtered at 5,000 Hz when paired with the word and 10,000 Hz for the other two conditions. The level of the broadband noise was 40 dB SPL.

Procedure. Participants performed a duration-matching task as in Experiment 1. In any given condition, either the word "do" or the drum strike was paired with the broadband noise. For each pair of stimuli, only the broadband noise changed in duration according to the participant's adjustment. The word "do" and drum strike did not change in duration. Participants made adjustments to the broadband noise duration by moving a slide switch throughout its range.

A slide on a response box controlled the comparison stimulus and after each match the range of adjustments was varied to ensure that participants were not focusing in on a single point on the slide (Schlauch et al., 2001). When the word "do" was the standard stimulus, the rule for selecting the duration range was identical to that for the 500 ms condition in Experiment 1. When the drum strike was the standard stimulus, the range of adjustment for the comparison broadband noise was 40 ms to a randomly selected 870–1480 ms. Prior to data collection, participants completed a rectangular versus rectangular condition. The standard was 50 ms for 10 matches and 500 ms for 10 matches. The order of the standard and comparison was counterbalanced.

Participants were randomly assigned to one of three groups. The only difference in procedure among the groups was the instructions they were given. In the first group (instruction set "null"), instructions contained no explicit information about including or ignoring certain aspects of the sounds they were hearing. They were simply instructed to adjust the duration of the broadband noise to be equal to that of the word "do" or the drum strike (Grassi & Darwin, 2001; Schlauch et al., 2001). The second group (instruction set "all") was instructed to include all aspects of the sounds when making their judgments. The third group (instruction set "ignore") was instructed to ignore the echo portion of sounds when making their duration judgments. From a technical standpoint, the word "do" was processed with reverberation despite that the participants were instructed using the term "echo." It was found during pilot data that "echo" was more accessible to naïve participants and was used in a manner that included the effects of reverberation (diffuse decay due to multiple room reflections) and echoes (reflected sound wave that is more delayed and distinguishable from the source). Instruction sets for the three groups are listed in the Appendix.

After a group assignment was made, participants were introduced to the experiment. First, participants were read the appropriate instruction set while viewing a written copy. Following the instructions, participants were familiarized with the task by matching the duration of two rectangular-gated broadband noises, one of which had a fixed duration of 500 ms. Participants completed 10 such matches: 5 with the standard stimulus presented first and 5 with the comparison stimulus presented first.

Each participant made 40 matches during the data collection phase of the experiment. Ten matches were completed for each of four standard stimuli to the rectangular-gated broadband noise comparison stimuli. Two of the standards were the drum strike and the word "do." The waveforms of these stimuli are shown in Figure 3. From Figure 3, it is clear that these are falling-intensity sounds. The other two standard stimuli were the same sounds shown in Figure 3 played backward making them rising-intensity sounds. These rising-intensity sounds were produced by simply presenting the digitized waveform in reverse order. For all conditions, the standard stimulus was presented first in half of the measurements, and the comparison stimulus was presented first in the remaining half of the measurements.

Results

Figure 4 illustrates the results obtained for duration-matching tasks for the 3 groups of participants administered different instruction sets. Above the labels on the abscissa for each stimulus in this figure are the data for the "null," "all," and "ignore" instruction sets, in order from left to right. In the familiarization condition, participants matched rectangular-gated noise to itself. As in Experiment 1, the group median matches for these control conditions were within a few percent of the standard value of 500 ms. This result demonstrates that participants could perform the task and that differences in instructions did not affect judgments of this stimulus. The difference between the comparison and standard stimulus at equal subjective duration shows also that rising-intensity natural sounds were judged longer than the broadband noise comparison stimulus.

Figure 5 illustrates the rising-falling duration ratio for the natural sounds for each of the instruction sets. This analysis effectively eliminates the broadband noise comparison stimulus and allows a direct comparison of the subjective durations for the rising-intensity and falling-intensity natural sounds. The "all" and "null" instruction sets yielded a pattern of results identical to those observed for the broadband noise stimuli in Experiment 1. It is clear from this figure that the instruction set affected judgments and within an instruction set participants treated the drum strike and the word "do" (labeled "speech" in the figure) similarly. For the "null" instruction set, falling-intensity sounds were judged shorter than rising-intensity sounds, a result consistent with other studies (Grassi & Darwin, 2001; Schlauch et al., 2001). For this "null" instruction-set condition, group-mean rising-falling duration ratios were 1.45 and 1.61 for the word "do" and the drum

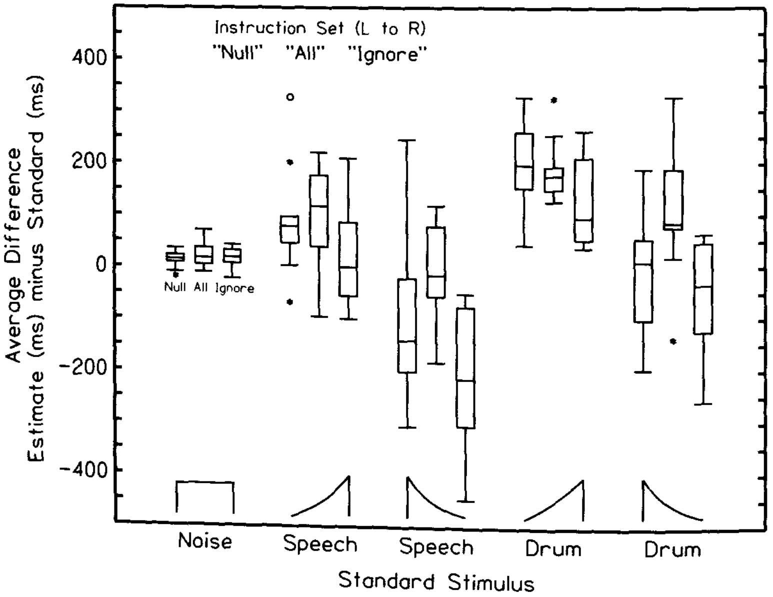

FIGURE 4 Box plot representations of the matching data for Experiment 2 for 10 participants. The data are represented as deviations of adjustments (ms) from the standard durations, which were 500 ms, 627 ms, and 435 ms for the rectangular-gated noise, speech (the word "do")., and the drum strike, respectively. The adjustable, comparison stimulus was a rectangular-gated broadband noise. Waveforms for the standard stimuli (rising intensity, falling intensity, or rectangular) for a particular condition are shown just above the abscissa. Above the labels on the abscissa for each stimulus are data for the "null," "all," and "ignore" instruction sets, in order from left to right.

strike, respectively. This finding is comparable to the rising-falling duration ratio of 1.4 (455 ms/315 ms) reported by Grassi & Darwin (2002). Grassi and Darwin (2001) used 500 ms, three-formant synthetic vowels with a raised cosine envelope to produce rising- and falling-intensity stimuli. There was no reverberation applied to their stimuli, however.

The instructions affected the perceived duration of both rising-intensity and falling-intensity stimuli. The greatest effect was an increase in the perceived duration for falling-intensity sounds in the "all" instruction set. This is illustrated in Figure 5, where the median rising-falling duration ratios at equal subjective duration for both stimuli are illustrated for the three instruction sets. The pattern of results is nearly identical for the drum strike and the word spoken in a simulated, reverberant environment. Specifically, when participants were told to include all aspects of the sound in their judgments, the rising-falling duration ratio was smaller than it was for the other two instruction sets. The instruction set to ignore echo portions increased the median rising-falling duration ratio for the word "do" compared to

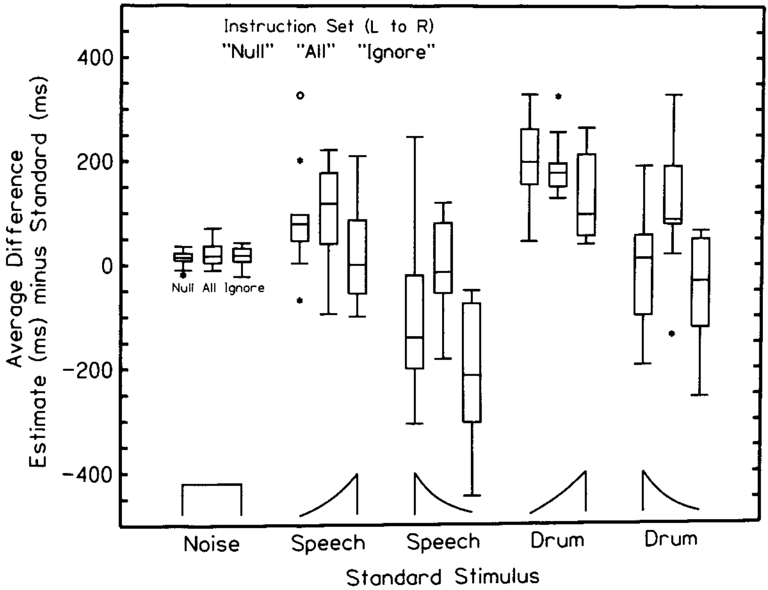

FIGURE 5 Box plot representations of the matching data for Experiment 2 for 10 participants. The data are plotted as rising-falling duration ratios for speech and drum sounds for 3 instruction sets. The data are represented in the same manner as in Figure 1.

the other conditions for the same stimulus, but individual variability was substantial for this instruction set. This same instruction, to ignore the echo, had a negligible effect on the rising-falling duration ratio for the drum strike compared to the ratio obtained with the "null" instruction set. The rising/falling duration ratio for equal subjective durations was examined in a repeated-measures ANOVA using stimulus and instruction set as the within-subject factors. There was a significant effect of instruction set, $F(2, 27) = 3.44; p < .05$. The effect of stimulus was not significant, $F(1, 27) = 0.02; p = .89$ and there was not a significant stimulus by instruction-set interaction, $F(2, 27) = 0.88; p = .42$. This analysis and the data shown in Figure 4 indicate that the rising-falling ratios were lower in the "all" instruction set than the "null" instruction set and the "ignore" instruction set was higher than the "null" instruction set.

Discussion

For the natural stimuli in this experiment, falling-intensity sounds were judged to be longer with the "all" instruction set than with the "ignore" instruction set. The "null" instruction set produced results nearly identical to those for the "ignore" instruction set. This pattern of results is consistent with the notion that a significant segment of the decay is normally ignored for falling-intensity sounds. Stecker & Hafter (2000) presented this explanation to account for differences in the subjective loudness and duration of rising-intensity and falling-intensity sounds. Accordingly, participants ignore the decay portions of falling-intensity sounds when there is reverberation because the tail, or some portion of the tail, of the decay is not considered a meaningful part of the sound being judged. Our result is consistent with the idea that participants adopt a criterion for judging when a rapidly decaying sound ends based on its sensation reaching a low but still audible level.

Rising-intensity sounds and falling-intensity sounds were judged to be longer when participants were given the "all" instruction set compared to when they received the other instruction sets ("null" and "ignore"). This demonstrates that this result is not unique to the artificially produced laboratory stimuli employed in Experiment 1. One might assume that participants' judgments were biased to some degree by the instructions. Despite a general lengthening of matched durations with the "all" instruction set, the rising-falling duration ratio became smaller, indicating that the effect of this instruction had a more pronounced influence on the falling-intensity sounds than rising-intensity sounds. If a simple judgment bias had been responsible, rising/falling duration ratios would have been unaffected across instruction sets because the bias would have increased the numerator and denominator comprising this ratio by equal proportions.

A notable difference between the results for Experiment 1 and Experiment 2 is the relation between the comparison duration for a rectangular-gated sound and the duration of the rising-intensity and falling-intensity sounds at equal-subjective duration. Prior matching studies have turned the same sound on and off in different manners to produce rising-intensity and falling-intensity sounds, as in Experiment 1. These studies have found that falling-intensity sounds represent an underestimation of the physical duration. Typically, rising-intensity sounds are judged to be about equal in duration to a rectangular-gated comparison stimulus (Schlauch et al., 2001). Experiment 2 is the first study to have participants match the durations of two different sounds. In this case, one was a natural sound and the other a broadband noise. It is not surprising that the broadband noise stimulus underestimated the duration of the natural stimuli for all conditions. Studies (Neuhoff, 1998, 2001; Schlauch et al., 2001) report that a broadband noise stimulus produces a smaller asymmetry in subjective duration for rising-intensity and falling-intensity sounds than the one observed for a natural stimulus. Consistent with this idea, the broadband noise comparison stimulus was adjusted to a longer duration than would be necessary for a natural comparison stimulus. The prominence of the natural stimuli is consistent with the localization priorities in a natural environment (Neuhoff, 1998). Broadband noise is associated with dispersed phenomena, such as wind and rain, and should not receive the same priority for preparation for source arrival as a stimulus more likely to be associated with a single sound source (Neuhoff, 2001).

EXPERIMENT 3: TEMPORAL MASKING PATTERNS FOR BROADBAND NOISES WITH RISING AND FALLING TEMPORAL ENVELOPES

The results of Experiment 1 showed that noise bands with rising temporal envelopes are perceived to be about 15% to 29% longer than sounds with falling temporal envelopes, even after participants are told to include all aspects of these sounds in their judgments. These figures are taken from the rising-falling duration ratios in the "all" instruction set for the 50 ms and 500 ms conditions, respectively. The asymmetry that remains after considering this natural bias might represent the contribution of a simple sensory process.

Sensory explanations for the asymmetries for sounds with rising and falling intensity have been identified in the peripheral and central auditory nervous system. A model of the auditory periphery encompassing auditory filtering and neural transduction of the auditory nerve (Patterson, Allerhand, & Giguère, 1995) responds asymmetrically to single cycles of rising- and falling-intensity sounds (Stecker & Hafter, 2000). Such predictions are in qualitative agreement with the loudness and subjective duration data for rising- and falling-intensity sounds for short duration sounds, but this model predicts frequency effects not observed in the data (Schlauch et al., 2001; Stecker & Hafter, 2000) and cannot account for observed asymmetries for sounds longer than 200 ms (Grassi & Darwin, 2001)

The possibility exists that the extent of the neural activity pattern in the auditory system may account for the differences in subjective duration between rising- and falling- intensity sounds that remain after participants are told to include all aspects of the sound. The temporal extent of neural excitation can be inferred from behavioral studies of non-simultaneous masking experiments where the masker and signal are not overlapped in time (Relkin & Turner, 1988; Turner, Relkin & Doucet, 1994).

The temporal masking pattern has proved important for modeling subjective duration. Fastl (1977) reported that subjective duration grows at a slower rate than physical duration for rectangular-gated noise and that this difference is consistent with the temporal masking patterns for these sounds. In his study, the subjective duration for a given stimulus was determined by observing times at which the temporal masking pattern dropped to a critical value close to detection threshold in quiet. These critical times mark the beginning and end of a sound for calculating subjective duration. To test this idea for rising- and falling-intensity sounds, we measured temporal masking patterns (the forward and backward masked sections) for the broadband noise stimuli in Experiment 1 and used these patterns to predict the rising- to falling- intensity differences in subjective duration using a method nearly identical to that of Fastl.

Method

Participants. Three adults with normal hearing participated in this study. Audiometric thresholds were 15 dB HL or better at octave frequencies between 250 Hz and 8000 Hz. Participants were paid for their participation.

Stimuli. The masker was a broadband noise with the rising- and falling-intensity characteristics described in Experiment 1. Masker duration was either 50 ms or 500 ms. The signal was a 10 ms, 2.0 kHz pure tone with Blackman ramps. A 10-ms signal was selected because it approximates the analysis window for temporal resolution (Viemeister & Plack, 1993). A 2.0 kHz tone was used as the signal frequency because (a) this frequency represents a low value on the minimum audibility curve and (b) thresholds in this frequency region have a smaller standard error of repeated estimates than ones obtained for frequencies of 4.0 kHz and higher. Participants listened monaurally through headphones (Telephonics, Model TDH-49P).

Procedure. Signal thresholds were assessed using a 3-alternative, forced-choice adaptive procedure that targeted 70.7% correct detections (Levitt, 1971). In this procedure, three sounds are presented in succession; all three contain the masker. A randomly selected interval also contains the signal. The participant was instructed to select the interval containing masker and signal. After two consecutive correct responses, the signal level was reduced. Any incorrect response resulted in an increase in the signal level. The initial step size was 3 dB and after two reversals the step size was reduced to 2 dB. The initial two reversals were not considered in the calculation of threshold. Threshold was calculated based on the average of the remaining reversals for an 80-trial block. For each condition, thresholds were based on a minimum of three blocks.

Masked thresholds for the tonal signal were measured for forward masking and backward masking. These thresholds were measured for the following masker-signal gaps: 5, 15, 25, 45, 85, and 165 ms. Gaps were calculated from the masker offset to signal center for forward masking and signal center to masker onset for backward masking. Detection thresholds for the tonal signal in quiet were also assessed. An exception to the above-mentioned conditions occurred for backward masking. In the backward-masking task, if a participant's threshold was within 5 dB of his or her detection threshold in quiet, threshold was only measured for two gap durations beyond the one that fell close to threshold. Prior to data collection, participants were given a minimum of two hours of practice.

Results

The group-mean results of the masking and detection tasks are shown in Figure 6. The upper panels illustrate results for the 50-ms conditions, and the lower panels

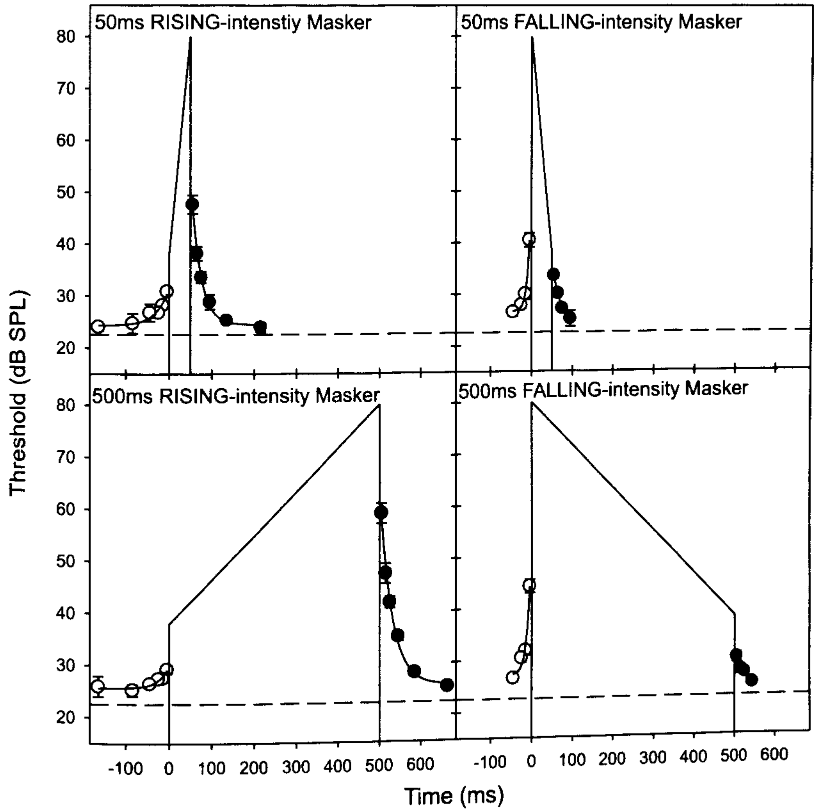

FIGURE 6 Forward and backward masked thresholds for rising- or falling intensity broadband noise maskers. Group-mean thresholds in the presence of 50-ms and 500-ms maskers are shown in the upper and lower panels, respectively. Open symbols represent backward-masked thresholds. Filled symbols represent forward-masked thresholds. The horizontal, dashed line represents the average 2.0 kHz threshold, in quiet, for the 3 participants. Error bars represent plus and minus one standard deviation.

illustrate results for the 500-ms conditions. Results for the rising- and falling-intensity maskers are shown in the left and right panels, respectively. As expected, the highest threshold elevation occurred when the masker was higher in level and the temporal proximity of the signal to the masker was small.

The temporal extent of these masking effects was quantified by fitting simple exponential decay curves to the threshold curves. Next, the times on these functions corresponding to signal thresholds within 5 dB, 10 dB, and 15 dB of detection threshold were found for each stimulus. The difference in time between these points plus the physical stimulus duration represents the prediction for subjective

TABLE 1
Predicted and Measured Rising-/Falling Ratio Comparison

Stimulus Duration	*Instruction Set "null" Matching*	*Instruction Set "all" Matching*	*Masking Pattern Criterion* 5 dB SL	*Masking Pattern Criterion* 10 dB SL	*Masking Pattern Criterion* 15 dB SL
50 ms	1.46 (0.18)	1.15 (0.10)	1.24	1.14	1.18
500 ms	1.85 (0.22)	1.29 (0.07)	1.12	1.12	1.09

Note. Rising-/falling-intensity ratios for duration-matching data for instruction sets "null" and "all" along with predicted ratios based on temporal masking patterns for three criteria for calculating the temporal extent of excitation. The predictions of the temporal masking patterns

duration (Fastl, 1977). If the amount of threshold elevation for a particular condition never exceeded the criterion values for 5, 10, or 15 dB SL, the nonsimultaneous masking for that condition did not contribute to the subjective duration. From these subjective duration estimates, a rising-falling duration ratio was calculated. Table 1 shows that all three criteria produced similar ratios.

Discussion

For both masker durations, the temporal extent of the masking pattern is greater for the rising-intensity masker than for the falling-intensity masker. This is a result of masker level effects and the asymmetry of forward and backward masking. Forward masking decays at a slower rate than backward masking for equal level maskers, and the asymmetry is greater for higher levels than for lower ones. Rising-intensity sounds terminate at a high level, which produces more forward masking than a falling-intensity sound that terminates at a low level. Even though a falling-intensity sound begins at a high level, the amount of backward masking produced by the falling-intensity masker does not compensate for the forward masking produced by the rising-intensity masker.

Predicted rising-falling duration ratios were calculated from the subjective durations estimated from the masking patterns. Table 1 shows that all three criteria (5 dB SL, 10 dB SL, and 15 dB SL) produced similar ratios all of which are greater than 1.0. This means the sound with the rising temporal envelope is predicted to have a longer subjective duration than a sound with a falling temporal envelope. These ratios are in qualitative agreement with the matching data, but the amount of the extension due to this masking effect is smaller than the observed difference in the matches. Further, the results for the instruction set "all," in which participants included all aspects of the sound in their judgment, are in closer agreement with the masking-pattern estimates than the "null" instruction set. This finding suggests that persistence may account for a portion of the asymmetry in subjective duration for rising-intensity and falling-intensity sounds.

GENERAL DISCUSSION

Sounds with falling-intensity envelopes are judged shorter than sounds with rising-intensity envelopes. The finding that the difference between perceived durations of rising-intensity and falling-intensity sounds is lessened by having participants attend to all aspects of these sounds supports Stecker & Hafter's (2000) hypothesis that listeners ignore part of the decay of a falling-intensity sound. This finding demonstrates that participants set a criterion for when they are willing to say that a sound ends, and this boundary is moveable. For sounds with falling-intensity envelopes, a sizeable portion of the sound is typically ignored. This behavior appears to be a general phenomenon for this class of sounds.

An instruction bias does not account for the entire magnitude of the asymmetry in subjective duration judgments for sound with rising and falling temporal envelopes. Even when participants considered all aspects of the stimulus, falling-intensity sounds were judged shorter than rising-intensity sounds by a small but significant amount. The source of this remaining difference may be related to a sensory process, possibly adaptation and persistence.

Persistence of perception can be thought of as the continuation of a sound's internal representation in the auditory nervous system after its source has stopped. Persistence is inferred often from psychoacoustic studies of masking obtained for non-simultaneously presented signals and maskers. An estimate of persistence can be obtained from the temporal masking pattern, which represents the effect of a masker on detection thresholds relative to the onset and offset of the masker. The temporal extent of the masking pattern for a sound is one method of inferring the internal representation of that sound at some unspecified level in the auditory system. Temporal-masking patterns obtained in this manner have proved important for modeling subjective duration (Fastl, 1977). The temporal masking patterns in experiment 3 are consistent with the idea that rising-intensity sounds persist longer than falling-intensity sounds and that this effect accounts for part of the subjective difference between these sounds.

As discussed, persistence may account only for part of the asymmetry in subjective duration observed between rising- and falling-intensity sounds. Adaptation is another possible mechanism that could make sounds with rising temporal envelopes more prominent than sounds with falling temporal envelopes. Adaptation is generally greater for high-level sounds than for low-level sounds. If the early, high-level portions of a sound with a falling temporal envelope were to adapt neurons responsible for coding the sound, latter portions would become inaudible or less audible. By contrast, the early, low-level segments of a sound with a rising temporal envelope would not cause substantial adaptation for latter, high-level segments of the sound. This idea, although consistent with the finding that sounds with falling temporal envelopes are perceived to be less loud and shorter than sounds with rising temporal envelopes, is inconsistent with the results of studies of loudness change (Neuhoff, 1998, 2001; Seifritz et al., 2002). Several studies have reported that sounds with rising envelopes change in loudness by a greater amount than sounds with falling envelopes (Neuhoff, 1998, 2001; Seifritz et al., 2002).

If adaptation were responsible for falling-intensity sounds being perceived as shorter than rising-intensity sounds because the final portions of the falling-intensity sound adapted to inaudibility, the loudness change would be greater for falling-intensity sounds than for rising-intensity sounds. Examples of these concepts and their consequences for perception are illustrated in Figure 7. The panels in this figure show hypothetical internal representations of loudness as a function of time for sounds with rising and falling temporal envelopes. For each of these panels, subjective duration is indicated by the horizontal extent of the shaded portion of the pattern, overall loudness by the area under the pattern, and loudness change (ΔL) by the difference in loudness at the beginning and end of the pattern. The lower panel illustrates the representation for a sound with a rising temporal envelope. Its duration and loudness are greater than those for the two upper panels representing the same sound reversed in time. Despite identical predictions for overall loudness and subjective duration, the prediction for loudness change differs in the upper and middle panels. The middle panel illustrates the hypothetical pattern showing the influence of adaptation. Adaptation resulted in the internal pattern terminating earlier than the physical duration of the stimulus. Overall loudness and subjective duration are reduced, but the loudness change (ΔL) is greater than it is for the sound that rises in intensity (lower panel). This finding, as noted earlier, is inconsistent with experimental results. By contrast, the top panel of this figure illustrates the internal representation for the sound with a falling temporal envelope that is not influenced by adaptation. In this example, the participant ignores the latter portion of the sound (the ignored segment indicated by dotted lines). This example yields predictions consistent with a variety of perceptual phenomena for these dynamically changing sounds. Overall loudness and subjective duration are reduced for this falling-intensity sound compared with the loudness for a rising-intensity sound (lower panel). Also, the loudness change (ΔL) is smaller for this falling-intensity sound than that for the sound with a rising temporal envelope that changes in loudness over its entire physical duration.

CONCLUSIONS

Rising-intensity sounds are perceived as longer than falling-intensity sounds and the instructions given to participants affect the magnitude of this difference. Participants have a natural bias to ignore a segment of the decay portion of falling-intensity sounds, a result consistent with a segregation resulting from a perceptual process. When participants are asked to include all aspects of sounds in their judgments, the difference in subjective duration for rising-intensity and falling-intensity sounds becomes considerably smaller. Even after considering the effect of instructions, rising-intensity sounds remain more prominent than

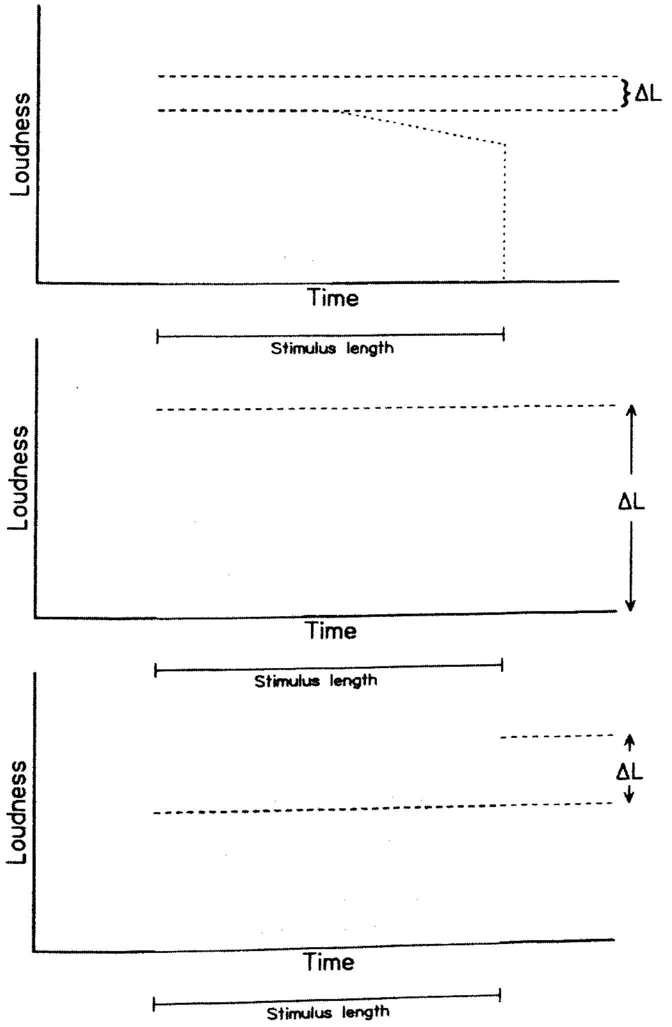

FIGURE 7 Schematic illustrations of possible internal representations for rising-intensity and falling-intensity sounds. The lower panel illustrates the internal representation for a rising-intensity sound. The middle and upper-panels illustrate the internal representations for falling-intensity sounds. The middle panel illustrates the possible effects of adaptation. The upper panel illustrates a condition in which the participant ignores the later portion of the sound (dotted line). Perceptual judgments are assumed to be based on shaded regions. Perceptual studies show a larger loudness change (ΔL) for the rising-intensity sounds than for falling-intensity sounds, a result consistent with the internal representation in the upper panel (ignoring the decay) and not middle panel (adaptation).

falling-intensity sounds (i.e., subjective durations for rising-intensity sounds are greater than for falling-intensity sounds). This remaining difference could be a result of an effect that stems from coding the stimulus at the auditory periphery—a sensory process that is not consistent with adaptation. One possible mechanism is related to persistence, which is consistent with the temporal masking patterns for these sounds.

REFERENCES

Abel, S. M. (1972). Duration discrimination of noise and tone bursts. *Journal of the Acoustical Society of America.*, *51*, 1219–1223.

ANSI S3.2–1995 (R2003). (1995). Bioacoustical terminology.

Barron, M., & Marshall, A. H. (1981). Spatial impression due to early lateral reflections in concert halls: The derivation of a physical measure. *Journal of Sound and Vibration*, 77(2), 211–232.

Benade, A. (1976). *Fundamentals of musical acoustics*. Oxford: Oxford University Press.

Bradbury, J. W., & Vehrencamp, S. L. (1998). *Principles of animal communication*. Sunderlan, MA: Sinauer.

Davis, H., & Silverman, S. R. (Eds.). (1978). Hearing and deafness (4th ed.). New York: Holt, Rinehart & Winston.

Fastl, J. (1977). Subjective duration and temporal masking patterns of broadband noise impulses. *Journal of the Acoustical Society of America*, *61*, 162–168.

Ghazanfar, A. A, Neuhoff, J. G., & Logothetis, N. K. (2002). Auditor looming perception in rhesus monkeys. *Proceedings of the National Academy of Sciences*, 99(24) 15755–15757.

Grassi, M., & Darwin, C. J. (2001). Perception of the duration of ramped and damped sounds with raised cosine amplitude. Fechner Day, October 22, 2001, *Proceedings of the Sixteenth Annual Meeting of the International Society for Psychophysics Germany*, October 19–23. 345–350.

Grey, J. M. (1977). Multidimensional perceptual scaling of musical timbres. *Journal of the Acoustical Society of America*, *61*, 1270–1277.

Hall, A. D., & Moore, D. R. (2003). Auditory neuroscience: The salience of looming sounds. *Current Biology*, *13*, R91–R93.

Levitt, H. (1971). Transformed up-down methods in psychoacoustics. *Journal of the Accoustics Society of America*, 49, 467–477.

Maier, J. X., Neuhoff, J. G., Logothetis, N. K., & Ghazanfar, A. A. (2004). Multisensory integration of looming signals by Rhesus Monkeys. *Neuron*, 43, 177–181.

Neuhoff, J. G. (1998). Perceptual bias for rising tones. *Nature*, *395*, 123–124.

Neuhoff, J. G. (2001). An adaptive bias in the perception of looming auditory motion. *Ecological Psychology*, *13*, 87–110.

Patterson, R. D. (1994a). The sound of a sinusoid: spectral models. *Journal of the Acoustical Society of America*, 96, 1409–1418.

Patterson, R. D. (1994b). The sound of a sinusoid: time-interval models. *Journal of the Acoustical Society of America*, 96, 1419–1428.

Patterson, R. D., Allerhand, M. H., & Giguère, C. (1995). Time-domain modeling of peripheral auditory processing: a modular architecture and software platform. *Journal of the Acoustical Society of America*, 98, 1890–1894.

Patterson, R. D., & Irino T. (1998). Modeling temporal asymmetry in the auditory system. *Journal of the Acoustical Society of America*, 104, 2967–2979.

Plomp, R. (1970). Timbre as a multidimensional attribute of complex tones. In R. Plomp and G. F. Smoorenburg, (Eds.), *Frequency analysis and periodicity detection in hearing* (pp. 397–414). Sijthoff, Leiden.

Relkin, E. M., & Turner, C. W. (1988). A reexamination of forward masking in the auditory nerve. *Journal of the Accoustical Society of America, 84,* 584–591.

Risset, J.-C., & Wessel, D. (1999). Exploration of timbre by analysis Synthesis. In D. Deutsch (Ed.), *The psychology of music* (pp. 113–169). San Diego: Academic Press.

Santana, C. (1999). Supernatural. New York: BMG/Arista.

Schlauch, R. S., Ries, D. T., & DiGiovanni, J. J. (2001). Duration discrimination and subjective duration for ramped and damped sounds. *Journal of the Acoustical Society of America, 109,* 2880–2887.

Schlauch, R. S., Ries, D. T., DiGiovanni, J. J., Elliot, S., & Campbell, S. (1998). Intensity discrimination of ramped and damped tones. *Journal of the Acoustical Society of America, 103,* 885–886.

Seifritz, E., Neuhoff, J. G., Bilecen, D., Scheffler, K., Mustoviic, H., Schachinger, H., et al. (2002). Neural processing of auditory looming in the human brain. *Current Biology, 12*(24), 2147–2151.

Stecker, G. C., & Hafter, E. R. (2000). An effect of temporal asymmetry on loudness. *Journal of the Acoustical Society of America, 107,* 3358–3368.

Tukey, J. W. (1977). *Exploratory data analysis.* Reading, MA: Addison-Wesley.

Turner, C. W., Relkin, E. M., & Doucet, J. (1994). Psychophysical and physiological forward masking studies: Probe duration and rise-time effects. *Journal of the Acoustical Society of America, 96,* 795–800.

Viemeister, N. F., & Plack, C. J. (1993). Time analysis. In W. Yost, A Popper ,& R. Fay (Eds.), *Human psychophysics, SHAR* (Vol. 3, pp. 116–154). New York: Springer Verlag.

Zahorik, P., & Wightman, F. L. (2001). Loudness constancy with varying sound source distance. *Nature Neuroscience, 4,* 78–83.

APPENDIX

The following written instructions were provided to participants in the duration matching experiments. All participants received Parts 1 and 3. Depending on the condition ("null," "all," or "ignore"), participants received a different Part 2 (the "null" instruction set did not contain a Part 2). The instructions in this appendix are numbered Parts 1, 2, and 3. The Part 2 sections are named ("null," "all," and "ignore"). Participants were given uninterrupted text containing only the instruction set relevant to their experimental condition.

Part 1. "You will hear two sounds played one after the other through the headphones. The sounds will repeat continuously, separated by brief silent periods. Your task is to match the duration of the two sounds as closely as possible. In order to match the duration of the two sounds you will move a slide on a response box that will increase or decrease the duration of one of the sounds."

Part 2. "Null": Nothing was added.

Part 2. "All": "You will notice that some or all of the sounds may have some soft, barely audible segments. For all of the sounds that you hear, include the entire sound in your judgment of its duration."

Part 2. "Ignore": "You will notice that some or all of the sounds may have an echo. When adjusting to the sounds that do have an echo, ignore the echo in your judgment of the sound's duration."

Part 3. "During a matching session, one sound will always remain fixed in duration while the other sound is adjustable. Please consider the whole range of adjustments available to you and bracket your adjustments several times around the point of equal duration before making your final decision. For example, move the slide to its upper and lower limits, and listen to the sounds at each of these points. Then move the slide several times above and below the point of equal duration, listening to the results of each adjustment. Finally, move the slide to the point where the duration of the two sounds is equal. When the slide is at the point of equal duration for both sounds, press any button on the response box for one second to save your result. The experimenter will then start a new sequence for you to complete."