Perception, 2012, volume 41, pages 594–605

doi:10.1068/p7128

Forward masking of dynamic acoustic intensity: Effects of intensity region and end-level

Kirk N Olsen, Catherine J Stevens
MARCS Institute, University of Western Sydney, Locked Bag 1797, Penrith, NSW 2751, Australia; e-mail: k.olsen@uws.edu.au; web: http://marcs.uws.edu.au
Received 23 September 2011, in revised form 7 April 2012

Abstract. Overestimation of loudness change typically occurs in response to up-ramp auditory stimuli (increasing intensity) relative to down-ramps (decreasing intensity) matched on frequency, duration, and end-level. In the experiment reported, forward masking is used to investigate a sensory component of up-ramp overestimation: persistence of excitation after stimulus presentation. White-noise and synthetic vowel 3.6 s up-ramp and down-ramp maskers were presented over two regions of intensity change (40–60 dB SPL, 60–80 dB SPL). Three participants detected 10 ms 1.5 kHz pure tone signals presented at masker-offset to signal-offset delays of 10, 20, 30, 50, 90, 170 ms. Masking magnitude was significantly greater in response to up-ramps compared with down-ramps for masker–signal delays up to and including 50 ms. When controlling for an end-level recency bias (40–60 dB SPL up-ramp vs 80–60 dB SPL down-ramp), the difference in masking magnitude between up-ramps and down-ramps was not significant at each masker–signal delay. Greater sensory persistence in response to up-ramps is argued to have minimal effect on perceptual overestimation of loudness change when response biases are controlled. An explanation based on sensory adaptation is discussed.

Keywords: adaptation, auditory looming, intensity change, loudness, persistence

1 Introduction

Acoustic intensity change is fundamental in real-world listening domains such as speech and music. In experiments, perception of intensity change can be investigated with continuously increasing (up-ramp) or decreasing (down-ramp) intensity envelopes. When these dynamic stimuli are presented with matched timbres (eg pure tones, square wave, synthetic vowel, or violin), frequency, duration, and range of intensity change, they are perceived very differently. For example, differences in perceived timbre have been reported. Brief (<100 ms) tonal up-ramps are perceived as sinusoidal, whereas temporally reversed down-ramps with identical energy spectra are perceived as hollow or percussive (Patterson 1994a, 1994b). Furthermore, the rise time of ramped stimuli affects perceived timbre. Tones with 10 ms up-ramp attacks tend to be perceived as 'plucked', whereas 70 ms up-ramp attacks tend to be perceived as 'bowed' (Cutting and Rosner 1974). Differences in subjective duration between up-ramps and down-ramps have also been reported. Tones and ecological sounds (eg the spoken word 'do' and a drum strike) $\leqslant 2000$ ms in duration are perceived longer in duration when presented as up-ramps, relative to down-ramps (DiGiovanni and Schlauch 2007; Grassi and Darwin 2006; Schlauch et al 2001). The 'echo hypothesis' offered by Grassi and Darwin explains this asymmetry with a decay-suppression mechanism, whereby the perceptual system 'ignores' the final decay of a down-ramp because of the decay's association with reverberation (see also Stecker and Hafter 2000).

On the dimension of loudness, up-ramps are perceived louder (Ries et al 2008; Stecker and Hafter 2000; Susini et al 2007) and judged to change more in loudness from onset to offset (Neuhoff 1998, 2001; Olsen et al 2010; Susini et al 2010). White-noise up-ramps do not typically elicit such perceptual overestimations (eg Neuhoff 1998). A greater perceived change in loudness in response to up-ramp tonal stimuli has been described as an adaptive survival response to an apparent looming or approaching sound source in the environment (Neuhoff 1998, 2001). The looming conjecture associates the continuous increase of intensity with an approaching object (Jenison 1997). Relative to static or receding objects, a perceptual bias will be selected for that affords extra time for a fight or flight response to an approaching and potentially threatening stimulus (see Pastore and Flint 2010 for a detailed discussion on the looming conjecture). This 'perceptual bias' in response to continuous increases of intensity change is argued to manifest as an overestimation of loudness change for up-ramps relative to down-ramps (Neuhoff 1998, 2001; but see Olsen and Stevens 2010). While an adaptive function suggests why perceptual overestimation of up-ramps may occur, a psychological explanation of the phenomenon awaits.

1.1 *Loudness change in response to dynamic acoustic intensity*

Overestimation of loudness change in response to up-ramps typically occurs when participants make judgments of an entire stimulus sweep (Neuhoff 1998; Olsen et al 2010; Susini et al 2010; Teghtsoonian et al 2005). When ratings of loudness change are calculated from static 'snapshots' or magnitude estimates of loudness at, for example, the beginning, midpoints, and end of long duration (10–180 s) dynamic stimuli, down-ramps are judged to change more in loudness than up-ramps when their offset level falls below 40 dB SPL (Canévet and Scharf 1990; Canévet et al 2003; Teghtsoonian et al 2000). This greater perceived-change of loudness for down-ramps has been termed 'decruitment' and is likely the result of intensity-region-specific sensory adaptation (Canévet and Scharf 1990; Canévet et al 1985; Scharf 1983). That is, the accelerated loss of loudness in response to down-ramps (decruitment) occurs as the end-point of the stimulus falls below 40 dB and approaches absolute threshold.

Intensity-region-specific effects on loudness change at shorter stimulus durations have been observed in response to up-ramp perception, with different mechanisms proposed. Teghtsoonian et al (2005) showed that global judgments of loudness change in response to up-ramps and down-ramps are dependent on the end-level intensity of pure tone and broadband up-ramps. For example, loudness change in response to 1.8 s linear 15 dB or 30 dB up-ramps increases as a function of stimulus end-level and not range of intensity change. That is, as the region of change (and thus the end-level) increases, so does the magnitude of loudness change. In global judgments of loudness change, an end-level bias could reflect memory-based recency, where responses are biased toward the most recent portion of the sound: the end-level (Susini et al 2007). However, research investigating the temporal weighting of global loudness using short (~1 s) duration stimuli comprising small intensity fluctuations elicited significant primacy effects, where earlier portions of a sound contribute to the overall loudness impression more so than mid- or end-points (eg Dittrich and Oberfeld 2009; Oberfeld and Plank 2011; Rennies and Verhey 2009). Small but significant recency effects in temporal weighting have also been reported (Dittrich and Oberfeld 2009; Pedersen and Ellermeier 2008).

Olsen et al (2010) investigated recency in memory and loudness change in response to dynamic intensity. From single and paired stimulus presentations, a significant overestimation of loudness change for speech (vowel) and musical (violin) 3.6 s up-ramps was observed in a balanced end-level design that controlled end-level differences (eg low region 50–70 dB SPL up-ramp versus high region 90–70 dB SPL down-ramp). In balanced end-level conditions, up-ramp overestimation of loudness change was eliminated in response to 1.8 s stimuli (cf Neuhoff 1998, 2001). However, as stimuli increase to 3.6 s, significant differences in loudness change remain and are not explained by the end-level bias. Persistence of neural excitation is an alternative sensory mechanism of differences in loudness change between up-ramps and down-ramps when post-stimulus perceptual judgments are made. That is, a longer post-stimulus *sensory* response of greater magnitude may result in a subjectively larger *perception* of change for that stimulus. This broad hypothesis is investigated in the present study.

1.2 *Sensory persistence, forward masking, and dynamic acoustic intensity*

The persistence of excitation explanation derives from the fact that the auditory system continues to respond to a sensory stimulus after it ceases to be presented. The principal behavioural method of investigating post-stimulus excitation is psychophysical forward masking. Forward masking is defined as an elevation of hearing threshold for a target signal presented after another stimulus event: the masker. The difference between masked signal threshold and signal threshold in quiet is an indicator of masking magnitude, a measure of the auditory system's response to a sensory stimulus beyond its physical presence at a particular point in time. Post-stimulus sensory response to short-duration (< 1 s) masking stimuli may last for durations up to ~200 ms but varies significantly as a function of masker intensity, frequency, and duration (eg Carlyon 1988; Elliott 1962, 1971; Fastl 1976a, 1976b, 1979; Fastl and Zwicker 2007; Houtgast 1972; Jesteadt et al 1982; Oxenham and Plack 2000; Plomp 1964; Ries et al 2008). In addition to neural persistence, adaptation is an alternative mechanism that may underlie forward masking, although models of temporal integration tentatively suggest that neural persistence provides a better description of forward masking data (eg Oxenham 2001).

Relative to down-ramps in the same region of intensity change, greater masking magnitude would be expected for up-ramp maskers (eg 60–80 dB SPL) when stimulus offset is 20 dB higher than down-ramps (eg 80–60 dB SPL). Therefore we hypothesise greater magnitude of masking in response to up-ramps in a high (60–80 dB SPL) and low (40–60 dB SPL) region of intensity change. However, the characteristics of masking in response to dynamic stimuli in a balanced end-level comparison remain to be investigated. Perhaps neural excitation in response to up-ramps is greater than that for down-ramps and is associated with up-ramp overestimation of loudness change in balanced end-level conditions reported elsewhere (eg Olsen et al 2010). However, if post-stimulus excitation is most closely associated with masker end-level, and not necessarily the global stimulus attributes from onset to offset, then no systematic differences in forward-masking magnitude will be observed.

Temporal masking patterns such as forward masking have been researched extensively. However, systematic investigations with masker envelopes that continuously increase or decrease in intensity are relatively rare (but see DiGiovanni and Schlauch 2007; Ries et al 2008) and no study to our knowledge reports a balanced end-level analysis using dynamic intensity maskers. In the majority of masking experiments intensity is treated as static by setting it to a particular steady-state level, ignoring the dynamic properties of sound that change through time and the perceptual biases that occur in response to increasing versus decreasing intensity. Ries et al (2008) provide an exception with temporal masking used to investigate backward, simultaneous, and forward-masking patterns of white-noise up-ramps and down-ramps. Three masker durations (10, 50, and 500 ms) and 0.5, 1.5, and 4.0 kHz 10 ms pure tone target signals were used to investigate masked thresholds. In forward-masking conditions, the pure tone signal was presented up to 145 ms beyond the offset of the masker stimulus and an exponential function was the best fit to describe the masked threshold decay. Although the main aim in Ries et al (2008) was to compare temporal masking data with predictions from dynamic loudness models (Chalupper and Fastl 2002; Glasberg and Moore 2002; see also Rennies et al 2010)—a different aim to the experiment reported here—the temporal masking paradigm employed by Ries et al (2008) used white-noise up-ramps and down-ramps and serves as a point of departure for the present study. That is, facets of the Ries et al (2008) temporal masking experiment have been incorporated here to investigate persistence of excitation in response to dynamic intensity stimuli, specifically in balanced end-level conditions. Consequently, the study presented white-noise maskers for methodological consistency with Ries et al (2008), and vowel maskers for comparison with loudness ratings in Olsen et al (2010) and Neuhoff (1998, 2001).

1.3 *Aim, design, and hypotheses*

The aim, using a forward-masking paradigm, was to investigate persistence of excitation in response to up-ramp and down-ramp stimuli, with specific consideration to the temporal masking characteristics of these stimuli with equivalent offset levels (ie a balanced end-level design). Masker stimuli were presented at durations of 3.6 s and the signal for detection was a 1.5 kHz 10 ms pure tone (cf Ries et al 2008). A $2 \times 2 \times 2 \times 6$ within-subjects design consisted of two masker timbres (white-noise, synthetic vowel) and two directions of masker intensity change (up-ramp, down-ramp) presented over two intensity regions (low 40–60 dB SPL, high 60–80 dB SPL). Two regions of change enabled a balanced end-level comparison in the analysis (low-region up-ramps compared with high-region down-ramps, both with end-levels of 60 dB). There were six masker-offset to signal-offset delays of 10, 20, 30, 50, 90, and 170 ms in each of the eight masker conditions. The dependent variable was the magnitude of masking (signal threshold elevation relative to signal threshold in quiet at six masker–signal delays).

For white-noise and vowel maskers, we hypothesise greater masking magnitude in response to up-ramps, relative to down-ramps, at each masker–signal delay. If persistence of excitation is most closely associated with the final portion of the masker (as is the case with steady-state maskers), then there will be no systematic differences in masking magnitude between up-ramps and down-ramps in a balanced end-level analysis (ie 40–60 dB SPL low-region up-ramps versus 80–60 dB SPL high-region down-ramps). However, it is conceivable that greater masking magnitude is recovered in response to up-ramps in the balanced end-level analysis, as the summation of energy in an up-ramp may lead to greater persistence response, relative to down-ramps that are characterised by a continuous fall in energy. This result would support persistence of excitation as a contributing mechanism associated with perceptual differences between balanced end-level up-ramps and down-ramps.

2 Method

2.1 *Participants*

Three normal-hearing participants, aged 29, 22, 25 years, completed the experiment: P1 (male) was the first author, P2 a female university employee, and P3 a male university graduate student. All binaural pure-tone thresholds were <18 dB HL at the 1.5 kHz audiometric frequency. Binaural thresholds in quiet for the 10 ms 1.5 kHz pure-tone used as the signal in the present experiment for P1, P2, and P3 were 17, 25, and 17.5 dB SPL, respectively. P1 and P3 had some prior experience in psychophysical experiments. All participants completed approximately 1 h of practice listening to various combinations of forward-masking conditions until they were comfortable with the procedure.

2.2 *Stimuli and equipment*

A 3.6 s synthetic vowel and a 3.6 s white-noise stimulus were used as maskers. Each masker consisted of either a linear intensity increase (up-ramp) or decrease (down-ramp) within the low (40–60 dB SPL) or high (60–80 dB SPL) region of intensity change. In total, there were eight different masker conditions. The generation of the vowel masker began with a 3.6 s steady-state synthetic vowel (/ə/) from a Klatt synthesiser (Klatt 1980) with the default sampling frequency of 8 kHz. Figure 1 displays the spectra of the vowel analysed in Praat (Version 5.3.06). The fundamental frequency of the vowel stimulus was 130.81 Hz. A 3.6 s steady-state white-noise was generated in Audacity (Version 1.3.3) with a sampling rate

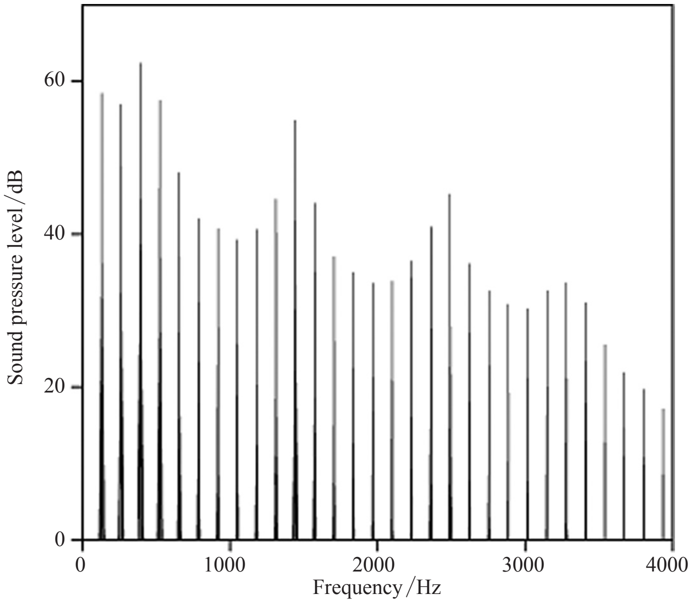

Figure 1. Spectra of the synthetic vowel (/ə/) analysed in Praat (Version 5.3.06).

of 8 kHz. 10 ms linear fade-in and fade-out ramps were included to remove onset/offset clicks. Up-ramps and down-ramps were constructed from the steady-state exemplars in a sound-attenuated booth using a custom computer program written in MAX-MSP (Version 4.6.3). First, the minimum and maximum level for both intensity regions was measured with a Brüel and Kjær Artificial Ear 4152 attached to a Brüel and Kjær hand-held Analyser 2250 using sound level meter Software BZ-7222. MAX-MSP generated an up-ramp and a down-ramp for vowel and white-noise maskers by using the minimum and maximum levels for each region as onset/offset anchors, and creating a linear intensity change between them with each original steady-state sound.

The signal used to measure thresholds in the presence of maskers was a 10 ms 1.5 kHz pure tone gated with 5 ms linear intensity (dB) rise/fall ramps. Masker–signal delays were measured in milliseconds from masker-offset to signal-offset (cf Fastl 1977; Oxenham 2001), resulting in six forward-masking signal delays: 10, 20, 30, 50, 90, and 170 ms. Measuring masker–signal delay from masker-offset to signal-offset in the 10 ms condition results in a signal onset in that condition of 0 s after the masker offset. In total, there were eight masker conditions comprising six masker–signal delays in each. The experiment protocol was custom written in Microsoft.Net (framework 3.5) using C# language and the experiment was conducted in a sound-attenuated booth. Stimuli were presented binaurally through Sennheiser HD 25 headphones.

2.3 *Procedure*

Signal thresholds were obtained using an adaptive three interval, three alternative forced-choice procedure targeting 70.7% of the psychometric function (Levitt 1971). In each trial, three intervals of the same masker condition were presented to participants, separated by 500 ms from offset to onset. A signal was randomly allocated to one interval at one of six masker–signal delays. Participants were instructed to select the interval that contained the masker and the signal by pressing 1 of 3 numerical keys on a computer keypad. A two-down one-up adaptive procedure was used to be consistent with other dynamic masking experiments (Ries et al 2008), where the signal intensity (dB SPL) decreased after two consecutive correct responses and increased after one incorrect response. The initial signal intensity began 10 dB below masker offset level. Initial step size was 5 dB for the first two reversals, then 2 dB for the remaining reversals. A run for each masker–signal delay consisted of thirteen reversals. The first three reversals were discarded; the levels of the remaining ten reversals were averaged to determine masked threshold for each masker–signal delay.

Participants completed eight sessions in total with each session conducted on a separate day. A session comprised one randomly chosen masker with six runs relating to the six masker–signal delay conditions. Three blocks were presented per session; each block contained two interleaved runs of randomly selected masker–signal delays for the session-specific masker. Participants were advised to rest for 10–15 min between blocks. A session lasted approximately 2 h, with all sessions taking approximately 16 h over 8 days.

3 Results

Figure 2 shows forward-masking thresholds averaged across participants and exponential functions for each curve of best fit. The use of exponential curves follows Ries et al (2008) where exponential decay was the best fit for forward-masking data in response to white-noise up-ramps and down-ramps. The mean difference in masking level (dB SPL) between up-ramps and down-ramps at each signal delay was calculated to investigate the characteristics of forward-masking magnitude between each dynamic stimulus (see table 1).

A full-factorial $2 \times 2 \times 2 \times 6$ repeated-measures analysis of variance (ANOVA) was conducted on forward-masking magnitude with Huynh–Feldt corrections to degrees of freedom. The ANOVA resulted in significant main effects of masker timbre ($F_{1,2} = 3772.54$, $p < 0.001$, Huynh–Feldt epsilon = 1.00), direction of intensity change ($F_{1,2} = 147.66$, $p < 0.01$, Huynh–Feldt epsilon = 1.00), and region of intensity change ($F_{1,2} = 643.73$, $p < 0.01$, Huynh–Feldt epsilon = 1.00). Masking magnitude was significantly greater for: (i) vowel maskers relative to white-noise maskers; (ii) up-ramps relative to down-ramps; and (iii) high regions of intensity change, relative to low regions. There was also a significant main effect of masker–signal delay ($F_{2.53,5.07} = 148.13$, $p < 0.001$, Huynh–Feldt epsilon = 0.51). Pairwise comparisons between the six masker–signal delays (Bonferroni adjusted α of 0.003) show that masking magnitude was greater at the 30 ms masker–signal delay relative to the 90 ms delay ($t_2 = 24.55$, $p = 0.002$), and greater at the 30 ms masker–signal delay relative to the 170 ms delay ($t_2 = 20.26$, $p = 0.002$). Masking magnitude was greater for the 50 ms masker–signal delay relative to the 90 ms delay ($t_2 = 49.62$, $p < 0.001$).

We hypothesised that up-ramps elicit greater magnitude of masking, relative to down-ramps, at each masker-signal delay. This hypothesis was supported in the main with a significant intensity direction × masker–signal delay interaction ($F_{5,10} = 59.33$, $p < 0.001$, Huynh–Feldt epsilon = 1.00). Using a Bonferroni adjusted α of 0.008, six pairwise comparisons were conducted between up-ramps and down-ramps at each masker–signal delay. Magnitude of masking was significantly greater for up-ramps relative to down-ramps at the 10 ms masker–signal delay ($t_2 = 30.38$, $p = 0.001$), 20 ms masker–signal delay ($t_2 = 14.73$, $p = 0.005$), 30 ms masker–signal delay ($t_2 = 17.06$, $p = 0.003$), and 50 ms masker–signal delay ($t_2 = 21.32$, $p = 0.002$). There were no significant differences between up-ramps and down-ramps at the 90 ms and 170 ms masker–signal delays (p values > 0.036).

3.1 *Balanced end-level planned contrasts*

Masking magnitude was also calculated under balanced end-level conditions. Specifically, planned contrasts were conducted comparing masking magnitude between 40–60 dB SPL low region up-ramps and 80–60 dB SPL high region down-ramps at each masker–signal delay. As can be seen graphically in figure 2 (lower panels) and numerically in table 1, there were no significant differences in masking magnitude between up-ramps and down-ramps with balanced end-levels of 60 dB (p values > 0.07).

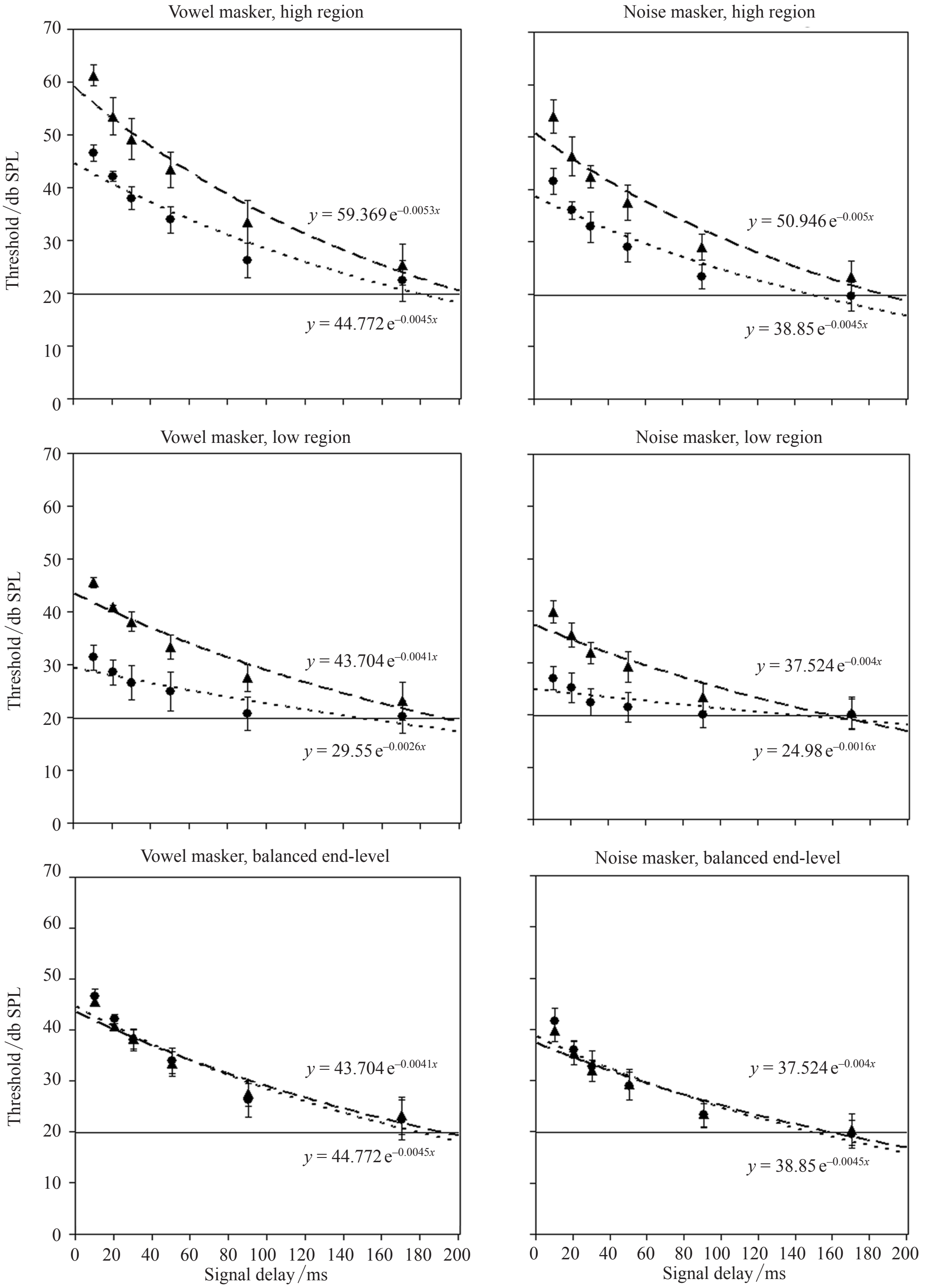

Figure 2. Mean forward-masking patterns ($N = 3$) from 3.6 s up-ramp (solid triangle) and down-ramp (solid circle) vowel and white-noise maskers in low (40–60 dB SPL) and high (60–80 dB SPL) regions of intensity change. Masker-offset to signal-offset delays were 10, 20, 30, 50, 90, and 170 ms. Balanced end-level comparisons (40–60 dB SPL up-ramps versus 80–60 dB SPL down-ramps) for vowel and white-noise maskers are also displayed. Exponential curves are shown in each panel for up-ramps (dashed line, top equation) and down-ramps (dotted line, bottom equation), and error bars represent standard error of the mean. Solid horizontal lines represent mean signal threshold in quiet.

Table 1. Difference between up-ramp and down-ramp mean forward-masking magnitudes.

Delay / ms	Vowel masker			Noise masker		
	low region	high region	balanced end-level	low region	high region	balanced end-level
10	14.20	14.73	−1.00	12.67	12.40	−1.80
20	12.27	11.40	−1.33	10.20	10.33	−0.60
30	11.60	11.23	0.16	9.47	9.60	−0.87
50	8.47	9.47	−0.60	7.80	8.47	0.33
90	6.73	7.33	1.27	3.33	5.80	0.20
170	3.03	3.07	0.80	0.33	3.80	0.93

Note. Table reports results of the difference of mean forward-masking magnitude (dB SPL) between up-ramp and down-ramp maskers for a 1.5 kHz 10 ms pure tone signal ($N = 3$). Positive numbers represent greater masking from up-ramps; negative numbers represent greater masking from down-ramps. Delay refers to signal delay (ms from masker-offset to signal-offset). Balanced end-level reports difference between 40–60 dB SPL (low region) up-ramps and 80–60 dB SPL (high region) down-ramps.

4 Discussion

Psychophysical forward masking in response to continuous increases (up-ramps) and decreases (down-ramps) of acoustic intensity was investigated. Past experiments report overestimation of loudness change in response to up-ramps relative to down-ramps matched on frequency, duration, range, and region of intensity change (eg Neuhoff 1998, 2001; Olsen et al 2010; Stecker and Hafter 2000; Susini et al 2010). In the present study we investigated the broad hypothesis that greater persistence of excitation in response to up-ramps is associated with perceptual differences in loudness change. Specifically, 3.6 s vowel and white-noise dynamic stimuli used in experiments investigating loudness (Ries et al 2008) and loudness change (Olsen et al 2010) were presented as maskers over two dynamic regions of intensity change (40–60 dB SPL, 60–80 dB SPL).

In both regions of intensity change, significantly greater magnitude of masking was obtained for up-ramps, relative to down-ramps, at masker-offset to signal-offset delays up to and including 50 ms. Typically, forward-masking magnitude increases as a function of intensity, even with short steady-state masker durations (eg 300 ms) (Jesteadt et al 1982). The key stimulus attribute determining magnitude of masking from the 3.6 s maskers in the present study may be the final portion of the masker (dominated by end-level) and not necessarily the relatively long 3.6 s intensity change preceding it. In earlier studies, greater end-levels for up-ramps relative to down-ramps in various regions of intensity change have made it difficult to disentangle effects of end-level from the direction of intensity change.

To control for differences in stimulus end-level, the influence of continuous intensity change in forward masking was investigated with a balanced end-level analysis, where up-ramps and down-ramps were compared between two regions of intensity change: a 40–60 dB SPL low region up-ramp compared with an 80–60 dB SPL high region down-ramp. If a greater response is associated with a greater perception of loudness change for up-ramps when end-level differences are controlled in this way (eg Olsen et al 2010), greater masking magnitude should be observed for up-ramps in the critical balanced end-level comparison shown in table 1 and illustrated in figure 2. However, this was not the case. There were no systematic differences in forward-masking magnitude between up-ramps and down-ramps with 60 dB end-levels. The balanced end-level analysis shows that post-stimulus persistence in response to dynamic auditory stimuli is primarily the result of end-level intensity and not the direction of intensity change preceding it. Therefore, differences in post-stimulus sensory excitation can be eliminated as a mechanism systematically contributing to perceptual differences in response to 3.6 s vowel and white-noise up-ramps and down-ramps when stimulus end-levels are equivalent.

Having ruled out differences in post-stimulus persistence, we are left with a yet-to-be-explained duration-specific perceptual phenomenon: continuous 3.6 s linear increases of acoustic intensity are perceived to change significantly more in loudness than continuous 3.6 s decreases when (a) each stimulus intensity change is of the same physical magnitude; and (b) each stimulus ends on an equivalent intensity (Olsen et al 2010). Sensory adaptation primarily affecting down-ramp perception at this duration is a likely explanation.

4.1 *Down-ramp adaptation and loudness change*

The effect of adaptation in response to continuous sweeps of decreasing intensity was originally discussed in decruitment experiments by using loudness ratio scales from magnitude estimation (Canévet and Scharf 1990; Canévet et al 2003; Schlauch 1992; Teghtsoonian et al 2000, 2005). Decruitment refers to a rapid decline in loudness in response to down-ramps of durations from 1 to 10 s, but most pronounced between 10 and 180 s. Decruitment manifests when the latter portions of a continuous decrease of intensity are perceived softer than those portions presented in isolation. For example, the end-level of a 20 s, 70–40 dB SPL continuous intensity sweep is perceived to be approximately a quarter of the loudness of the same 40 dB end-level presented as a short isolated tone burst (Teghtsoonian et al 2005). Simple and induced loudness adaptation has been proposed as a contributing mechanism (Canévet and Scharf 1990; Canévet et al 1985; Scharf 1983).

The effect of adaptation in response to down-ramps has been neglected in studies focusing on overestimation of loudness change in response to up-ramps (eg Neuhoff 1998, 2001; Olsen et al 2010). These particular experiments use interval rather than ratio loudness scaling. Interval scales such as the visual analogue scale used in Neuhoff (1998, 2001) and Olsen et al (2010) do not determine the physical difference in decibels between up-ramps and down-ramps; they tell us only that there is a perceived difference. From interval scales it is equivocal whether the perceptual phenomenon is up-ramp overestimation or down-ramp underestimation. Indeed, the rapid decline of loudness as the continuous intensity sweep of a down-ramp approaches its end-level may be associated with the yet-to-be explained perceptual differences between up-ramps and down-ramps under balanced end-level conditions (Olsen et al 2010).

Until now, a balanced end-level analysis has been defined purely in physical terms. In the present study, forward masking between up-ramps and down-ramps with balanced end-levels of 60 dB were compared. In Olsen et al (2010), loudness change was investigated between up-ramps and down-ramps with balanced end-levels of 70 dB. Consideration of a balanced end-level analysis in perceptual terms, and not just physical terms, has yet to be discussed. As Canévet and Scharf (1990) alluded to, the previous higher intensity portion of a down-ramp at any temporal point (except its onset) may adapt the subsequent portion and result in a perceptually 'softer' stimulus that accumulates until offset. An up-ramp ending on the same physical level as the down-ramp would not theoretically elicit this adaptation effect, as each portion of its intensity sweep is preceded by a lower intensity. In this example, the end-level of the down-ramp is perceived to be softer than the end-level of the up-ramp; hence a physical balanced end-level analysis reflects unbalanced end-level perception, where differences in loudness change can be interpreted as down-ramp underestimation. The predominant theoretical focus on an adaptive, survival response to looming auditory motion (Neuhoff 1998, 2001) has resulted in the interpretation of differences in loudness change in favour of up-ramp overestimation, but we see here that this is not necessarily the case.

Down-ramp *underestimation* cannot be ruled out. A related explanation based on down-ramp underestimation can also be found in the echo hypothesis and related decay-suppression mechanism (Grassi and Darwin 2006; Stecker and Hafter 2000). The decay of a down-ramp is associated with reverberation in the environment. The final portion of the decay may be perceptually 'ignored' and removed from the computation of subjective judgments, resulting in down-ramp underestimation.

Future studies should not only analyse perceived differences between up-ramps and down-ramps with balanced physical end-levels, but also differences in *physical* intensity between up-ramps and down-ramps when end-levels are *perceptually* equivalent. Measuring the point of subjective equality (Pastore and Flint 2010) of dynamic and steady-state stimuli will show whether differences in loudness derive from up-ramp overestimation or down-ramp underestimation, and the stimulus conditions under which this phenomenon systematically varies. Additional work on this topic could also compare recent loudness models (eg Chalupper and Fastl 2002; Glasberg and Moore 2002; Meddis and O'Mard 2005; Plack et al 2002). Such models may not only predict differences in forward masking between up-ramps and down-ramps, but also the potential role of masking within a masker (eg the extent to which the first 100 ms of the masker masks the following 100 ms, and so on).

Finally, it is important to demarcate experiments investigating dynamic looming motion from experiments investigating loudness change in response to dynamic acoustic intensity. The salience of real and apparent looming auditory/visual motion is not in doubt (eg Ball and Tronick 1971; Freiberg et al 2001; Ghazanfar et al 2002; Hall and Moore 2003; Maier et al 2008; Rosenblum et al 1993; Schiff and Oldak 1990). It is invoking an adaptive survival function in the context of psychophysical investigations of loudness change in response to up-ramps that is premature.

Acknowledgments. This research was supported by an Australian Postgraduate Award held by the first author and an Australian Research Council Discovery Project grant (DP0771890) held by the second author. We thank Johnson Chen for computer programming, and Matt Petoe and John Cass for helpful advice.

References

Ball W, Tronick E, 1971 "Infant responses to impending collision: Optical and real" *Science* **171** 818

Canévet G, Scharf B, 1990 "The loudness of sounds that increase and decrease continuously in level" *Journal of the Acoustical Society of America* **88** 2136–2142

Canévet G, Scharf B, Botte M-C, 1985 "Simple and induced loudness adaptation" *International Journal of Audiology* **24** 430–436

Canévet G, Teghtsoonian R, Teghtsoonian M, 2003 "A comparison of loudness change in signals that continuously rise or fall in amplitude" *Acta Acustica united with Acustica* **89** 339–345

Carlyon R P, 1988 "The development and decline of forward masking" *Hearing Research* **32** 65–79

Chalupper J, Fastl H, 2002 "Dynamic loudness model (DLM) for normal and hearing-impaired listeners" *Acta Acustica united with Acustica* **88** 378–386

Cutting J E, Rosner B S, 1974 "Categories and boundaries in speech and music" *Attention, Perception, & Psychophysics* **16** 564–570

DiGiovanni J J, Schlauch R S, 2007 "Mechanisms responsible for differences in perceived duration for rising-intensity and falling-intensity sounds" *Ecological Psychology* **19** 239–264

Dittrich K, Oberfeld D, 2009 "A comparison of the temporal weighting of annoyance and loudness" *Journal of the Acoustical Society of America* **126** 3168–3178

Elliott L, 1971 "Backward and forward masking" *International Journal of Audiology* **10** 65–76

Elliott L L, 1962 "Backward and forward masking of probe tones of different frequencies" *Journal of the Acoustical Society of America* **34** 1116–1117

Fastl H, 1976a "Temporal masking effects: I. Broadband noise masker" *Acustica* **35** 287–302

Fastl H, 1976b "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" *Journal of the Acoustical Society of America* **61** 162–168
Fastl H, 1979 "Temporal masking effects: III. Pure tone masker" *Acustica* **43** 282–294
Fastl H, Zwicker E, 2007 *Psychoacoustics: Facts and Models* (Berlin: Springer)
Freiberg K, Tually K, Crassini B, 2001 "Use of an auditory looming task to test infants' sensitivity to sound pressure level as an auditory distance cue" *British Journal of Developmental Psychology* **19** 1–10
Ghazanfar A A, Neuhoff J G, Logothetis N K, 2002 "Auditory looming perception in rhesus monkeys" *Proceedings of the National Academy of Sciences of the USA* **99** 15755–15757
Glasberg B R, Moore B C J, 2002 "A model of loudness applicable to time-varying sounds" *Journal of the Audio Engineering Society* **50** 331–342
Grassi M, Darwin C J, 2006 "The subjective duration of ramped and damped sounds" *Perception & Psychophysics* **68** 1382–1392
Hall D A, Moore D R, 2003 "Auditory neuroscience: The salience of looming sounds" *Current Biology* **13** R91–R93
Houtgast T, 1972 "Psychophysical evidence for lateral inhibition in hearing" *Journal of the Acoustical Society of America* **51** 1885–1894
Jenison R L, 1997 "On acoustic information for motion" *Ecological Psychology* **9** 131–151
Jesteadt W, Bacon S P, Lehman J R, 1982 "Forward masking as a function of frequency, masker level, and signal delay" *Journal of the Acoustical Society of America* **71** 950–962
Klatt D H, 1980 "Software for a cascade/parallel formant synthesizer" *Journal of the Acoustical Society of America* **67** 971–995
Levitt H, 1971 "Transformed up-down methods in psychoacoustics" *Journal of the Acoustical Society of America* **49** 467–477
Maier J X, Chandrasekaran C, Ghazanfar A A, 2008 "Integration of bimodal looming signals through neuronal coherence in the temporal lobe" *Current Biology* **18** 963–968
Meddis R, O'Mard L P, 2005 "A computer model of the auditory-nerve response to forward-masking stimuli" *Journal of the Acoustical Society of America* **117** 3787–3798
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
Oberfeld D, Plank T, 2011 "The temporal weighting of loudness: effects of the level profile" *Attention, Perception, & Psychophysics* **73** 189–208
Olsen K N, Stevens C J, 2010 "Perceptual overestimation of rising intensity: Is stimulus continuity necessary?" *Perception* **39** 695–704
Olsen K N, Stevens C J, Tardieu J, 2010 "Loudness change in response to dynamic acoustic intensity" *Journal of Experimental Psychology: Human Perception and Performance* **36** 1631–1644
Oxenham A J, 2001 "Forward masking: Adaptation or integration?" *Journal of the Acoustical Society of America* **109** 732–741
Oxenham A J, Plack C J, 2000 "Effects of masker frequency and duration in forward masking: Further evidence for the influence of peripheral nonlinearity" *Hearing Research* **150** 258–266
Pastore R E, Flint J, 2010 "Magnitude judgments of loudness change for discrete, dynamic, and hybrid stimuli" *Attention, Perception, & Psychophysics* **73** 886–907
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
Pedersen B, Ellermeier W, 2008 "Temporal weights in the level discrimination of time-varying sounds" *Journal of the Acoustical Society of America* **123** 963–972
Plack C J, Oxenham A J, Drga V, 2002 "Linear and nonlinear processes in temporal masking" *Acta Acustica united with Acustica* **88** 348–358
Plomp R, 1964 "Rate of decay of auditory sensation" *Journal of the Acoustical Society of America* **36** 277–282

Rennies J, Verhey J L, 2009 "Temporal weighting in loudness of broadband and narrowband signals" *Journal of the Acoustical Society of America* **126** 951–954
Rennies J, Verhey J L, Fastl H, 2010 "Comparison of loudness models for time-varying sounds" *Acta Acustica united with Acustica* **96** 383–396
Ries D T, Schlauch R S, DiGiovanni J J, 2008 "The role of temporal-masking patterns in the determination of subjective duration and loudness for ramped and damped sounds" *Journal of the Acoustical Society of America* **124** 3772–3783
Rosenblum L D, Wuestefeld A P, Saldana H M, 1993 "Auditory looming perception: Influences on anticipatory judgments" *Perception* **22** 1467–1482
Scharf B, 1983 "Loudness adaptation", in *Hearing: Research and Theory* Eds J V Tobias, E D Schubert (New York: Academic Press) pp 1–56
Schiff W, Oldak R, 1990 "Accuracy of judging time to arrival: Effects of modality, trajectory, and gender" *Journal of Experimental Psychology: Human Perception and Performance* **16** 303–316
Schlauch R S, 1992 "A cognitive influence on the loudness of tones that change continuously in level" *Journal of the Acoustical Society of America* **92** 758–765
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
Stecker G C, Hafter E R, 2000 "An effect of temporal asymmetry on loudness" *Journal of the Acoustical Society of America* **107** 3358–3368
Susini P, McAdams S, Smith B K, 2007 "Loudness asymmetries for tones with increasing and decreasing levels using continuous and global ratings" *Acta Acustica united with Acustica* **93** 623–631
Susini P, Meunier S, Trapeau R, Chatron J, 2010 "End level bias on direct loudness ratings of increasing sounds" *Journal of the Acoustical Society of America* **128** EL163–EL168
Teghtsoonian R, Teghtsoonian M, Canévet G, 2000 "The perception of waning signals: Decruitment in loudness and perceived size" *Perception & Psychophysics* **62** 637–646
Teghtsoonian R, Teghtsoonian M, Canévet G, 2005 "Sweep-induced acceleration in loudness change and the 'bias for rising intensities' " *Perception & Psychophysics* **67** 699–711