Journal of Experimental Psychology:
Human Perception and Performance
2009, Vol. 35, No. 6, 1898–1912

0096-1523/09/$12.00 DOI: 10.1037/a0015018

Changes in the Perceived Duration of a Narrowband Sound Induced by a Preceding Stimulus

Robert P. Carlyon, John M. Deeks, Yury Shtyrov, Jessica Grahn, Hedwig E. Gockel, Olaf Hauk, and Friedemann Pulvermüller
Medical Research Council Cognition and Brain Sciences Unit

The authors show that a narrowband noise (NBN) is perceived as longer when presented immediately after a wideband noise (WBN), compared to when the WBN is absent. This effect depended on the WBN's frequency spectrum overlapping that of the NBN, and it increased as the duration of the WBN increased up to 300 ms. It decreased when a silent gap was introduced between the WBN and NBN, but remained significant for an easily detectable gap of 40 ms. A correlate of the effect was observed in the mismatch negativity (MMN) to a deviant stimulus, consisting of a WBN + NBN, presented in a sequence of more common isolated WBNs. The MMN latency was longer for an on-frequency than for an off-frequency WBN; and, more importantly, the size of this difference correlated across participants with the difference in perceived duration. A rhythm-adjustment experiment showed that the presence of an on-frequency WBN immediately preceding a tone caused that tone to be heard as starting earlier than when the WBN was absent. The results are discussed in relation to the continuity illusion and models of duration encoding.

Keywords: continuity illusion, duration, mismatch negativity, rhythm

In everyday life, people often encounter situations where the temporal structure of a sound is ambiguous due to the potentially masking effects of other stimuli. For example, if a portion of one sound is briefly masked by another, then the auditory system must "decide" whether that sound remained on during the masker or whether it was interrupted. An important and well-studied phenomenon is that a perception of continuity can occur even when a sound is physically turned off for a short time, provided that the resulting silent interval is filled with another, *inducing* stimulus (Houtgast, 1972; Vicario, 1960; Warren, Obuseck, & Ackroff, 1972). An example of this *continuity illusion* is shown in Figure 1A, in which a portion of a tone is replaced by a noise. Decades of research have produced a comprehensive body of knowledge concerning the stimulus parameters that govern the continuity illusion: For example, a useful rule of thumb is that the inducing sound would have masked the interrupted sound if it had really remained on, and the illusion is weakened or absent when there are clear silent gaps between the interrupted and inducing sounds (Bregman, 1990; Warren, 1999). It is also worth noting that the illusion depends on the portion of the stimulus occurring after, as well as before, the noise (Ciocca & Bregman, 1987).

Here we report and investigate another illusory phenomenon that depends on ambiguity of the temporal boundaries of a sound. An example of the stimulus is shown in the left-hand part of Figure 1B and consists of a wideband noise (WBN) immediately followed by a narrowband stimulus that can be a tone or, in most of our experiments, a narrowband noise (NBN). It resembles the continuity illusion stimulus shown in Figure 1A, except that the narrowband sound occurs only after the WBN. As we show, the perceived duration of this NBN is several tens of milliseconds longer than that of an NBN presented in isolation (right part of Figure 1B). We argue that although the WBN causes the NBN to be heard as longer rather than causing it to sound continuous, this *duration illusion* has several factors in common with, and is likely to arise from similar mechanisms to, the continuity illusion. As with the continuity illusion, it depends on the frequency relationship between the WBN and NBN and is reduced or absent when there is a silent gap between the wideband and narrowband sounds (cf. right-hand side of Figure 1A). Broadly speaking, the duration illusion could arise in one of two ways, each of which has different implications for the way sounds are encoded by the auditory system.

One class of explanation states that the NBN is heard to start earlier than it would have if presented in isolation and before the end of the WBN. If the time at which it is perceived to end is not affected by the WBN, then this could lead to a perceived duration increase. This would be evidence that perception lags the stimulus by a variable amount, because we do not normally hear the end of an isolated WBN as containing an NBN. Furthermore, the size of this lag would have to be at least as long as the increase in the perceived duration, so measures of the duration increase would impose a lower limit on the lag between stimulus and perception. This in turn could rule out explanations for the illusion based on stages of processing where the latency of the neural response is shorter than the perceived duration increase. It is also worth noting that this class of explanation is roughly consistent with models of duration encoding in which the brain counts the number of inter-

---

Robert P. Carlyon, John M. Deeks, Yury Shtyrov, Jessica Grahn, Hedwig E. Gockel, Olaf Hauk, and Friedemann Pulvermüller, Medical Research Council Cognition and Brain Sciences Unit, Cambridge, England.

Correspondence should be addressed to Robert P. Carlyon, Medical Research Council Cognition and Brain Sciences Unit, 15 Chaucer Road, Cambridge CB2 7EF, England. E-mail: bob.carlyon@mrc-cbu.cam.ac.uk

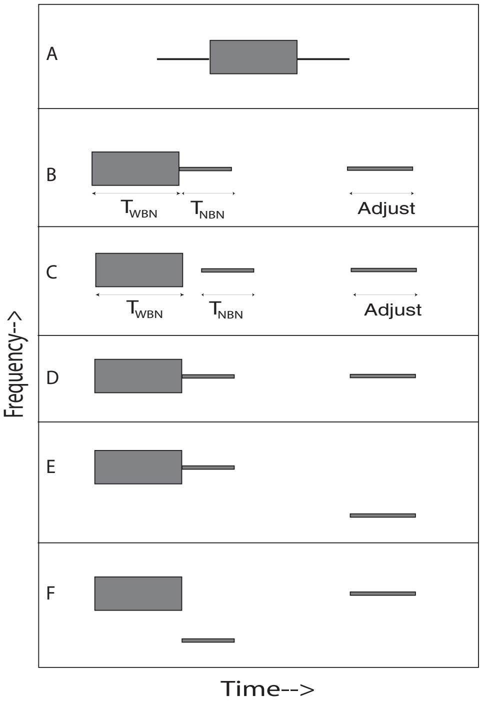

Figure 1. Schematic spectrograms of some of the stimuli used in the experiments described here and elsewhere. T_{WBN} = duration of wideband noise; T_{NBN} = duration of narrowband noise.

nally generated pulses between the start and end of a stimulus (for a review, see Allan, 1979). Such "counting" models would have to be modified, however, so that the estimate of the start of the NBN was affected by the WBN. One relevant modification to such models has recently been proposed by Tsuzaki and colleagues (Tsuzaki & Kato, 2000; Tsuzaki & Tanaka, 2003) and is considered in the Discussion section.

Alternatively, it may be that our perception of duration is not simply encoded as the time elapsed between the perceived start and end of a stimulus. Evidence that duration can be recoded comes from neural recordings from the auditory midbrains of mice and from the auditory cortex of cats (Brand, Urban, & Grothe, 2000; He, Hashikawa, Ojima, & Kinouchi, 1997; Xia, Qi, & Shen, 2000), which have revealed neurons "tuned" to different durations. It is possible that the WBN affects the processing of duration at a stage of processing where explicit information on onset and offset time is not retained—for example, by differentially affecting the responses of duration-sensitive cells. If so, then it could cause the NBN to be heard as longer but not to start earlier. This would be an example of sounds being represented in terms of auditory features that are not necessarily consistent. Evidence that the auditory system operates this way comes from a previous study of the continuity illusion in our laboratory: Listeners perceived a sinusoidally frequency modulated (SFM) tone as continuing behind another SFM tone having the same carrier frequency but a different modulation rate (Carlyon, Micheyl, Deeks, & Moore, 2004). Under such circumstances, listeners hear two tones that have the same carrier frequency but different modulator rates, even though, if both modulations were physically present, listeners would hear a complex pattern of modulation (Carlyon et al., 2004; Lyzenga, Carlyon, & Moore, 2005).

Here we investigate the increase in perceived duration using a combination of behavioral and electrophysiological techniques. First, we obtained behavioral measures of the illusion as a function of the duration of the WBN, of the NBN, and of the effect of introducing a silent gap between them. Combined with the results of a gap-detection experiment, we show that the illusion persists (albeit in a reduced form) even when there is a clearly detectable silent interval between the two stimuli. A further experiment was performed to confirm that the effects observed were due to a genuine increase in the perceived duration of the NBN that immediately follows the WBN, rather than to a decrease in the perceived duration of the subsequent isolated comparison NBN (cf. Scharf, Buus, & Nieder, 2002). We then report an EEG study measuring the mismatch negativity (MMN) to an occasional NBN that occurs in a sequence of WBNs. Our results show that the MMN to the NBN occurs later in a condition where the illusion occurs, compared to one where it does not. Importantly, this increased latency correlated, across listeners, with the increase in perceived duration. This result suggests that the MMN reflects some additional processing that occurs when the illusion is generated, but that, at the stage where the MMN is generated, an increase in perceived NBN duration is not accompanied by an earlier response to its onset. In contrast, our final experiment, which involves perceptual judgments of rhythm, does suggest that at this stage of processing, the NBN is judged as starting earlier when the illusion occurs compared to when it does not.

Experiment 1: Effects of NBN Duration

Rationale and Method

In this experiment we introduce our basic technique for measuring the illusion and study the effects of the NBN duration. Each trial consisted of a test stimulus followed 500 ms later by a comparison stimulus. In the wideband condition, the test stimulus consisted of a 500-ms WBN immediately followed by an NBN. To generate the WBN, we first constructed a 4-s noise by summing equal-amplitude components from 1000 to 4000 Hz (0.25-Hz spacing) in random phase. Its overall level was 78 dB SPL, and its spectrum level was 43 dB SPL. On each trial a 500-ms portion of this file was selected at random and was turned on and off with 3-ms raised cosine ramps. The NBN was constructed in a similar manner, contained components between 1900 and 2100 Hz, and had the same spectrum level as the WBN. It was also turned on and off with 3-ms raised-cosine ramps, and the start of its onset ramp coincided with the end of the offset ramp of the WBN at the zero-voltage points. Its duration, T_{NBN}, was 50, 100, 200, 300, 400, 500, or 700 ms, in different conditions.

The comparison stimulus consisted of an NBN that was drawn from the same file, but at a different start point, as used for the NBN in the test stimulus. At the start of each run its duration was selected at random from a rectangular distribution spanning the range 0.5 to 1.5 times the duration of the NBN in the test stimulus. At the end of each trial the participant could adjust the duration of the comparison stimulus to be presented on the next trial by pressing one of four buttons on a response box; these could increase or decrease the duration by a factor of 1.41 or 1.1. The participant was instructed to match the duration of the comparison stimulus to the NBN in the test stimulus, to ignore the WBN, and to bracket the perceived durations when performing the match. When satisfied with the match, the participant pressed a fifth button and could then initiate the next run by pressing a sixth *start* button.

Two additional conditions were included as controls. To check for any response bias in the adjustment procedure, we made the silent condition the same as the wideband condition, except that the WBN was replaced by 500 ms of silence. To check for any general influence of turning on a sound before the test NBN, the WBN was replaced by a notched noise in the notch condition. The notched noise contained components between 365–1461 Hz and 2730–4634 Hz and had the same spectrum level and overall bandwidth as the WBN. Its loudness was therefore similar to that of the WBN.

In this experiment, as well as in Experiments 2–4, all stimuli were played out with 16-bit resolution and a sampling rate of 40 kHz, using a CED 1401plus laboratory interface (Cambridge Electronic Design, Cambridge, England). They were then lowpass filtered at 9200 Hz (Kemo VBF5.01; Kemo Ltd., Beckenham, United Kingdom; attenuation rate 100 dB/octave), attenuated (TDT PA4; Tucker-Davis Technologies, Alachua, FL), and played via a homemade headphone amplifier to one earpiece of a Sennheiser HD250 headset (Sennheiser Electronic Corp., Old Lyme, CT). All stimuli were calibrated using a B&K 4153 artificial ear (Brüel & Kjær, Nærum, Denmark) and an HP3561A dynamic signal analyzer (Agilent Technologies, Wokingham, United Kingdom).

Four normal-hearing listeners, all experienced in psychoacoustical experiments, took part. One was John M. Deeks, and the others were paid for their participation. Each sat individually in a double-walled sound-insulating room. Each testing session typically lasted for 2 hr, with breaks as needed. All conditions were run in an interleaved random order, and the mean of between 5 and 15 runs was used to represent each condition. Throughout this article, geometric rather than arithmetic means were calculated across runs and across participants.

Results

The dashed lines in Figure 2 show the matched duration of the NBN, minus the veridical duration, in the three conditions of Experiment 1. The first four panels show data for individual participants, with mean data shown in the last panel. Although the pattern of results varies somewhat across participants, two trends are apparent in the data for all. First, the matched duration is generally longer in the on-frequency condition (open upward triangles) than in either the notch (inverted triangles) or silence conditions (circles). Second, the illusory increase in matched duration generally increases with increases in the veridical duration of the NBN. These trends were confirmed by a two-way (Condition × Duration) repeated-measures analysis of variance (ANOVA), which revealed main effects of condition, $F(2, 6) = 13.5$, $p < .02$, $\eta^2 = .818$; and of duration, $F(6, 18) = 3231.5$, $p < .001$, $\eta^2 = .999$. The interaction of condition with duration was not significant.[1]

Averaged across participants, the illusory increase in duration grew from about 16 ms at $T_{NBN} = 50$ ms to about 136 ms at $T_{NBN} = 700$ ms. Data from the silence condition (circles) suggest that very little of this effect reflected a response bias: The largest value obtained was about 19 ms at $T_{NBN} = 500$ ms. Somewhat larger values, reaching 56 ms at $T_{NBN} = 500$ ms, were obtained in the notch condition (downward triangles). This could have been due to some nonspecific effect of having any sound presented before the NBN or, alternatively, to energy leaking into the auditory filters responding to the NBN, thereby leading to an attenuated version of the illusion. To check the amount of this leakage, we passed our notched and wideband noises through the auditory filter model proposed by Glasberg and Moore (1990), and found that the response of a filter centered on 2000 Hz was 26.4 dB lower for the notched than for the wideband noise.

The solid line with filled diamonds in Figure 2 shows the difference in matched duration between the wideband and notched conditions. This value, which can be considered to represent a conservative estimate of the increase in perceived duration, is roughly constant at about 50 ms for T_{NBN} between 200 and 500 ms and then increases to 113 ms for $T_{NBN} = 700$ ms. Part of this increase, however, may be due to the decrease in the matched duration in the notched condition at $T_{NBN} = 700$ ms, which was due to the results of participant P3.

Experiment 2: Effect of WBN Duration

Rationale and Method

In Experiment 2 we studied the effect of the WBN duration on the illusion, using a larger number of participants, and replaced the notched condition with one using a noise that produced less excitation in auditory filters responding to the NBN.

The stimuli and methods were similar to those of Experiment 1, but T_{NBN} was fixed at 300 ms, and the effect of increasing the WBN duration (T_{WBN}) from 0.1 to 1.0 s was studied. In addition to providing more information on the maximum size of the illusion, we wished to distinguish between two ways in which the illusion could depend on T_{WBN}. If the illusion were driven by an aftereffect of the onset of the WBN, then, as T_{WBN} is increased beyond some value, this would increase the time between the NBN and WBN onsets, causing the illusion to decrease. Alternatively, if the illusion were driven predominantly by the portion of the WBN immediately preceding the WBN, then the increase in perceived duration could increase monotonically up to some value of T_{WBN} and then reach an asymptote.

Apart from varying T_{WBN} instead of T_{NBN}, the methods used were identical to those of Experiment 1, with two exceptions. First, to use a control sound that was less likely than the notched noise of Experiment 1 to excite auditory filters responding to the NBN, we replaced the notched noise with an off-frequency noise that contained frequency components between 5000 and 8000 Hz. It was generated in the same way as the (on-frequency) wideband noise and had the same level. The silence condition was omitted. Second, the number of participants was increased to 8; three of these, all of whom were paid volunteers, had participated in

---

[1] In this article we adjust significance levels using the Huynh-Feldt sphericity correction but report the uncorrected degrees of freedom. All data were log-transformed prior to analysis, but similar trends were obtained throughout when untransformed data were analyzed.

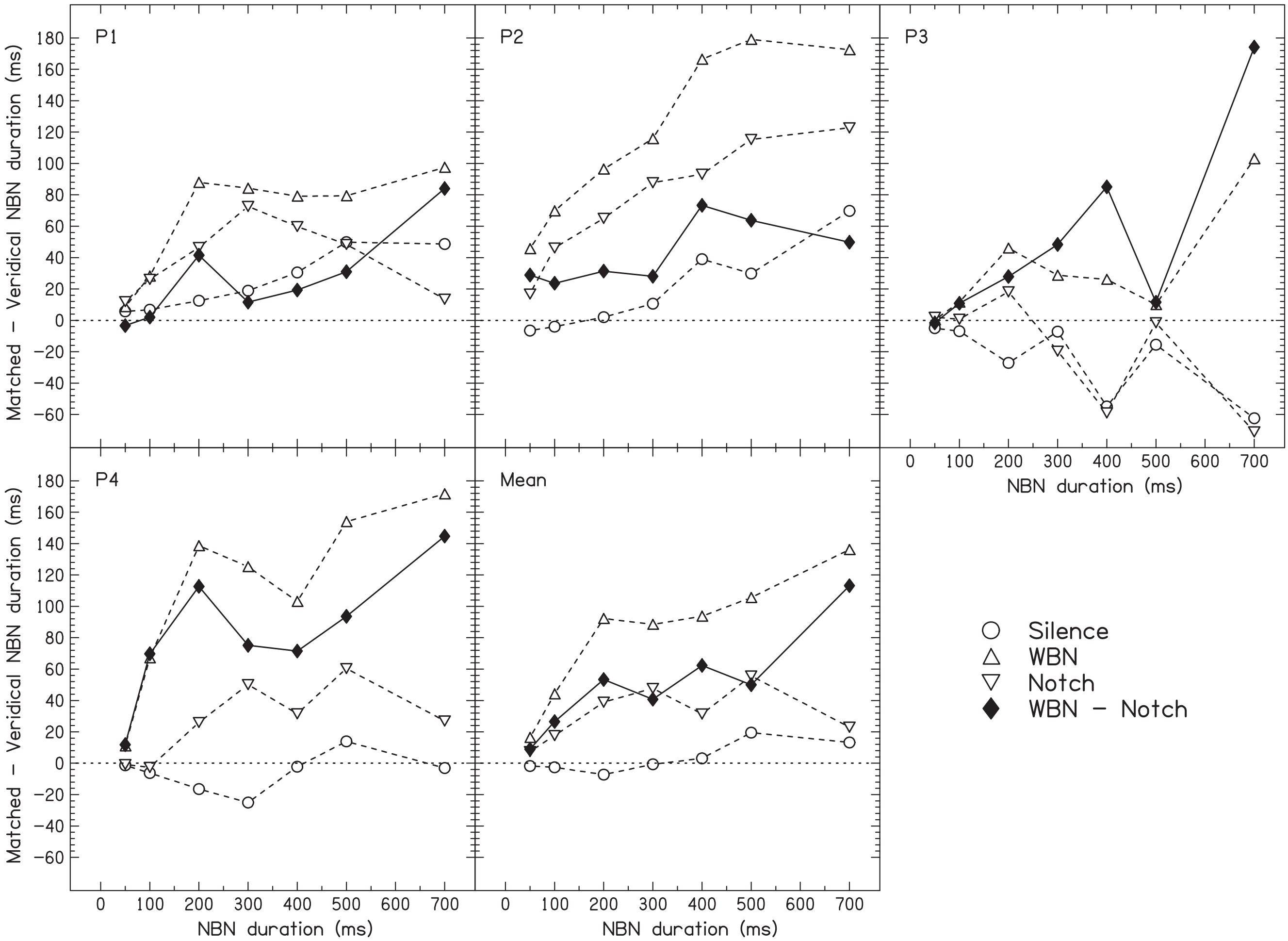

Figure 2. Results of Experiment 1 for individual participants (P1, P2, P3, P4; first four panels) and averaged across participants (last panel). Each curve joining open symbols shows the difference between the adjusted and veridical duration of the narrowband noise (NBN) as a function of the veridical NBN duration, in one condition. The line connecting the solid diamonds shows the difference between the matches obtained when the NBN was preceded by an on-frequency wideband noise and those obtained when the NBN was preceded by a notched noise.

Experiment 1 (see Table 1). Data were averaged from between 8 and 15 runs per participant, depending on the time available for testing.

Results

The results of Experiment 2 are shown in Figure 3, averaged across participants and with standard errors shown by error bars. For the on-frequency WBN (upright triangles), the matched duration exceeded the veridical NBN duration of 300 ms by an amount that increased from 21 ms at T_{WBN} = 100 ms to about 67 ms at T_{WBN} = 300 ms. A one-way ANOVA performed on the results for this condition alone revealed a highly significant effect of T_{WBN}, $F(6, 42) = 10.7$, $p < .001$, $\eta^2 = .603$. The duration increase was significantly larger than with an off-frequency WBN (inverted triangles), as evidenced by a main effect of condition in a two-way (Condition × T_{WBN}) repeated-measures ANOVA, $F(1, 7) = 189$, $p < .001$, $\eta^2 = .964$. That ANOVA also revealed a main effect of T_{WBN}, $F(6, 42) = 10.9$, $p < .001$, $\eta^2 = .610$, but the interaction was not significant. The difference between the results obtained in the on- and off-frequency conditions is shown by the filled triangles.

It is important to note that the increase in perceived duration in the on-frequency condition did not change as T_{WBN} was increased beyond 300 ms: Planned comparisons revealed that although the effect at T_{WBN} = 300 ms was significantly greater than at each of the two shorter durations, it did not differ significantly from that at larger values of T_{WBN}. This finding has implications for one explanation for the illusion, discussed in the next section.

One further finding is worth remarking on. The results for the 3 participants who also took part in Experiment 1 were similar in the two experiments for the condition that they had in common (T_{WBN} = 500 ms, T_{NBN} = 300 ms). In both experiments the effect differed substantially across the 3 listeners (P1, P2, P3), being 83, 113, and 23 ms in Experiment 1 and 60, 128, and 30 ms in

Table 1
List of Experiments in Which Each Participant Took Part

Participant	Experiment 1	2	3	4	5	6
P1	1	2	3	4	5	6
P2	1	2				
P3	1	2				
P4	1					
P5		2				
P6		2				
P7		2				
P8		2	3	4		
P9		2	3			
P10			3			
P11				***4***		***6***
P12				4		
P13				4	5	6
P14				***4***	***5***	
P15				4		6
P16				4		6
P17					5	
P18					5	
P19					5	
P20					5	
P21					5	6
P22					5	
P23					5	
P24					5	
P25					5	
P26					5	
P27					5	
P28					5	
P29						***6***
P30						6
P31					5	6

Note. The three participants who did not show the duration illusion are indicated by bold italics.

Experiment 2. Hence both the average size of the effect and its variation across participants seem stable and largely unaffected by modest changes in experimental context. A list of the experiments performed by all participants described in this article is shown in Table 1.

Experiment 3: Effect of Gap Between WBN and NBN

Rationale

It is likely that the auditory system has a less accurate estimate of the time an NBN starts when it is immediately preceded by a WBN than when it is presented in isolation. Hence the increase in perceived duration could occur as a result of the system "guessing" the NBN's onset time as somewhere between the start and end of the WBN. This would be consistent with the increase in the size of the illusion as T_{WBN} is increased from 100 to 300 ms, because the time range over which the NBN could be estimated to start increases with T_{WBN}. One would, however, have to posit some limitations on this guess to account for the fact that the perceived duration of the NBN does not increase further as T_{WBN} is lengthened beyond about 300 ms.

The effect of T_{WBN} could alternatively, or additionally, be due to the WBN masking the onset of the NBN and the fact that forward masking increases as a function of masker duration (Carlyon, 1988; Zwislocki, Piroda, & Rubin, 1959). To test this general idea, Experiment 3 measured the effects of introducing a silent gap (Δt) between the WBN and the NBN (Figure 1C). First, Experiment 3A measured the increase in perceived duration at two values of T_{WBN} as a function of Δt. Second, Experiment 3B measured the smallest gap that could be detected for a range of values of T_{WBN}, including those used in Experiment 3A.

Method

Experiment 3A used a similar method to Experiment 2, the only differences being that only two values of T_{WBN}, 200 and 500 ms, were used, and that the silent gap between the WBN and NBN could not only be 0 ms, as in Experiment 2, but alternatively, 5, 10, 20, 40, or 80 ms. Conditions were run in a random order, and the average of between 9 and 17 runs (depending on the time availability of the participant) was used to represent each condition. As in Experiment 2, the measures were performed using both an on-frequency (1000–4000 Hz) and an off-frequency (5000–8000 Hz) WBN. Also as in that experiment, all noises were turned on and off with 3-ms raised-cosine ramps. Four participants took part.

Experiment 3B measured gap detection thresholds using a two-interval, two-alternative forced choice method with the same participants as in Experiment 3A. One interval of each trial contained a WBN followed by an NBN with $\Delta t = 0$ ms, whereas in the other interval the value of Δt was larger. The two intervals were separated by 300 ms. In different conditions the value of T_{WBN} was set to 100, 200, 300, 400, or 500 ms. At the start of each run, Δt was set to a large value (typically 100 ms). After each trial, the participant indicated which interval contained the nonzero gap and received correct-answer feedback. Initially, the value of Δt was reduced by a factor of 1.414 after every two consecutive correct answers and increased by the same factor after every incorrect

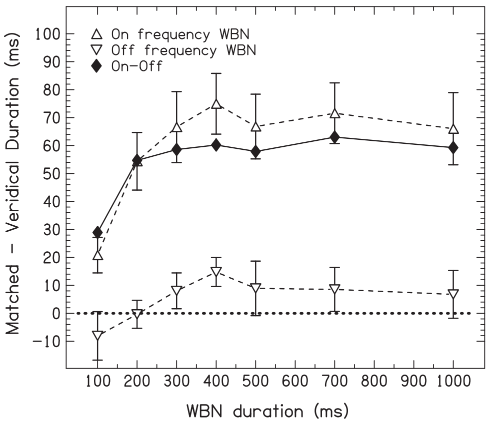

Figure 3. Results of Experiment 2, averaged across participants, with standard errors indicated by error bars. Each curve joining open symbols shows the difference between the adjusted and veridical duration of the narrowband noise (NBN) as a function of the duration of the on-frequency (upright triangles) or off-frequency (inverted triangles) wideband noise (WBN). The line connecting the solid diamonds shows the difference between the matches obtained when the NBN was preceded by an on-frequency WBN and those obtained when the NBN was preceded by an off-frequency WBN.

answer. The change from increasing to decreasing Δt or vice versa defined a turnpoint, and after 4 turnpoints, the factor was reduced to 1.189. Each run ended when 16 turns had occurred. The estimate for each run was obtained from the mean of the last 12 turnpoints, and each threshold here was estimated from the mean of at least five runs per condition, except when the maximum gap duration of 200 ms was exceeded. This occurred on four occasions for 1 participant, and when it did, only completed runs were included.

Results

The results of Experiment 3A are shown for the 200-ms and 500-ms WBN durations in Figures 4A and 4B, respectively. The dashed lines in each panel show the difference between the adjusted and actual NBN duration for individual participants, and the heavy solid line shows the mean data. At both values of T_{WBN}, the difference between adjusted and actual NBN duration decreases as a function of Δt. The value of Δt up to which this difference differs significantly from zero varied somewhat across participants: At T_{WBN} = 500 ms it was 80 ms (the longest value tested) for P9 and P10, 40 ms for P1, and 20 ms for P8. Averaged across participants, the adjusted duration differed significantly from the veridical value of 300 ms at Δt up to 20 ms at T_{WBN} = 200 ms, and up to 40 ms at T_{WBN} = 500 ms. Although this effect appears in the mean data to be larger at the longer WBN duration, this difference just failed to reach significance in a two-way repeated-measures ANOVA, $F(1, 3) = 7.9$, $p = .07$, $\eta^2 = .725$. That ANOVA confirmed the large effect of Δt apparent in the data, $F(5, 15) = 36.3$, $p = .001$, $\eta^2 = .924$; the interaction between T_{WBN} and Δt was not significant.

The gap detection thresholds from Experiment 3B are shown as a function of T_{WBN} for the on-frequency WBN in Figure 4C. Thresholds were between 14 and 20 ms for all values of T_{WBN} and did not vary significantly across this range. It is of interest that these thresholds were generally lower than the longest value of Δt at which the matched duration of the NBN in Experiment 3A exceeded its veridical value. This is particularly apparent in the data of P9 and P10; for T_{WBN} = 500 ms, they had a mean gap detection threshold of 19 ms, but the WBN continued to affect the matched duration of the NBN for gaps up to at least 80 ms. In addition, it is worth noting that thresholds were independent of T_{WBN}. This argues against the idea that the increase in the duration illusion with increasing T_{WBN} is due to an increase in the forward

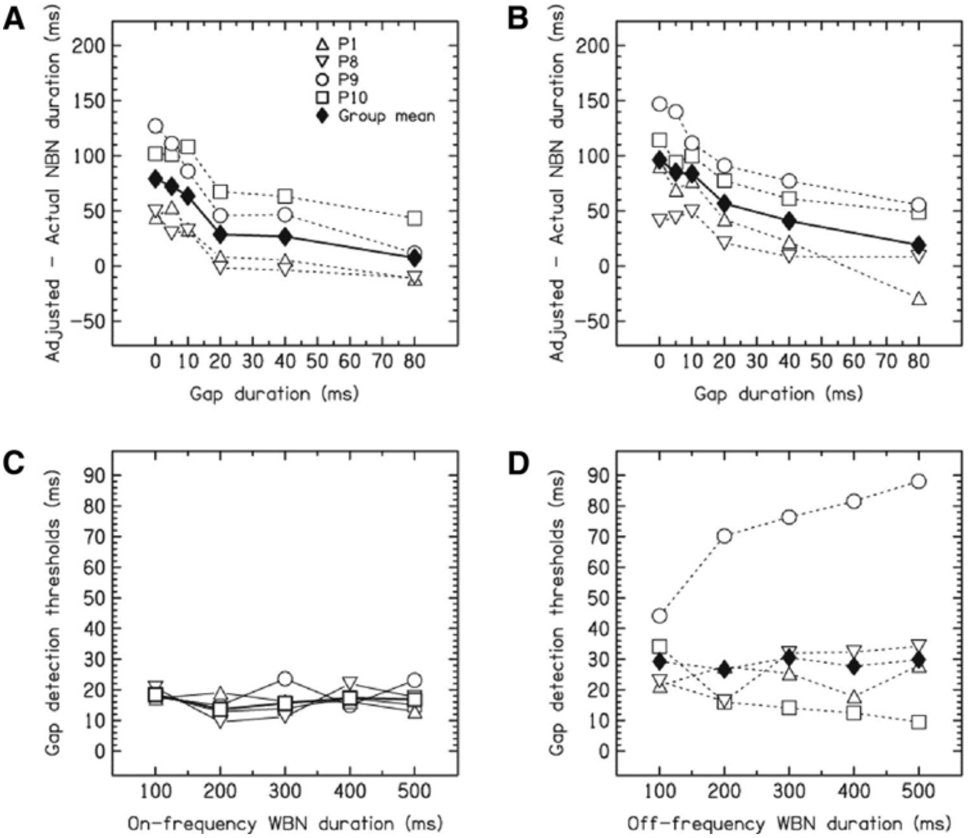

Figure 4. Panels A and B show the results of Experiment 3A for wideband noise (WBN) durations of 200 and 500 ms, respectively. Each curve joining open symbols shows the difference between the adjusted and veridical narrowband noise (NBN) duration as a function of the duration of a silent gap between the WBN and NBN, for 1 participant. The curves joining solid symbols show the data averaged across participants (P1, P8, P9, P10). Panels C and D show the gap-detection thresholds obtained with an on-frequency and an off-frequency WBN, respectively.

masking of the NBN onset with increasing WBN duration (Carlyon, 1988; Zwislocki, Piroda, & Rubin, 1959).

Figure 4D shows the gap detection thresholds obtained with the off-frequency WBN. Generally, these thresholds appear higher than with the on-frequency WBN, which would be consistent with evidence that listeners are relatively insensitive to timing differences between different frequency regions (Broadbent & Ladefoged, 1959; Vliegen, Moore, & Oxenham, 1999). This could occur even though the frequency difference between WBN and NBN may have allowed a better estimate of the absolute onset time of the NBN; the poor performance may have been due to listeners hearing a clear NBN onset in both intervals of each forced-choice trial. However, we should note that a two-way ANOVA revealed that the effect of WBN type failed to reach significance, $F(1, 3) = 4.2$, $p = .13$, $\eta^2 = .581$.

To summarize the results with an on-frequency WBN, the Δt at which the perceived duration increase disappears is, on average, longer (20–40 ms) than the smallest detectable gap of around 14–20 ms. For individual subjects, this discrepancy can be somewhat larger. Hence, although the data are consistent with the illusion being dependent on the ease with which the NBN onset can be detected (as is also the case for the continuity illusion), some duration increase can occur under conditions where the gap is detectable.

Experiment 4: Control for Effect of WBN on the Perceived Duration of the Comparison Stimulus

Rationale and Method

Experiments 1 to 3A required participants to adjust the duration of an isolated comparison NBN so that its perceived duration matched that of the test NBN. The results showed that when the WBN preceding the test has a spectrum encompassing that of the test NBN, the duration to which the comparison sound is adjusted exceeds that of the test. We have interpreted this finding as evidence that the WBN *increases* the perceived duration of the *test* NBN. However, another explanation is logically possible: namely, that the WBN *decreases* the perceived duration of the comparison NBN. Although this may at first mention seem implausible, a related phenomenon has been observed in the loudness domain. It has been shown that when an intense sound precedes a target, the intensity to which a subsequent comparison sound is adjusted in order to match the loudness of the test is greater than when the intense inducing sound is absent. Scharf et al. (2002) have argued that this finding is not due to the inducer increasing the perceived loudness of the test but rather reflects a long-term effect that reduces the loudness of the later comparison sound. The test sound is assumed to be largely immune to this effect because the influence of the inducer on the loudness of a subsequent sound increases over the first 1–2 s after the inducer ends.

Experiment 4 tested whether the WBN affected the perceived duration of the test or of the comparison NBN. The rationale was that the effect of a WBN on a subsequent NBN should be greatest when the frequency content of the WBN encompassed that of the NBN. There were three main conditions, named according to the frequency content of the WBN, test NBN, and comparison NBN; *H* stands for high frequency and *L* for low frequency. In condition HHH (Figure 1D), all noises were geometrically centered on 1000 Hz: The WBN contained frequencies from 707 to 1414 Hz, and the test and comparison NBNs spanned the range from 900 to 1100 Hz. We expected that, as in Experiments 1–3, the comparison NBN would be adjusted to a duration longer than that of the test NBN, regardless of which of these two sounds was affected by the WBN. The HHL condition (Figure 1E) was similar, except, crucially, the comparison NBN had a lower frequency (450–550 Hz). Any effect of the WBN on the perceived duration of the comparison sound should be reduced in this condition. In contrast, in condition HLH (Figure 1F), the test NBN spanned 450–550 Hz, and the comparison NBN spanned 900–1100 Hz. Here, any effect of the WBN on the perceived duration of the comparison sound should be the same as in condition HHH, but its influence on the perceived duration of the test should be reduced. To control for any response biases associated with adjusting NBNs of different frequencies to have the same duration, all three conditions were compared to control conditions in which the WBN was absent (SHH, SHL, and SLH, respectively, where S stands for silence). We therefore report the adjusted duration of the comparison NBN in each condition with the adjusted duration in the appropriate control condition subtracted out (e.g., HHH − SHH, HHL − SHL, and HLH − SLH). The prediction is that if the WBN affects the comparison, then the adjusted duration should be longer for the HLH than for the HHL condition—that is, HLH − SLH > (HHL − SHL), whereas if the WBN affects the test then the opposite should be true.

Stimuli were played out of an SB Live! sound card at a sampling rate of 44100 Hz, low-pass filtered at 20000 Hz (Kemo VBF/01), attenuated (TDT PA4), and presented via a headphone amplifier to one earpiece of a Sennheiser HD250 headset. In all other respects, the methods were similar to those in Experiments 1–3A. The overall level of the WBN was 57 dB SPL, and that of the high- and low-frequency NBNs was 50 dB SPL. As in previous experiments, the spectrum level of the WBN was the same as that of the (high-frequency) NBN that fell into the frequency region encompassed by it and had a value of 27 dB SPL. The WBN and NBN durations were 300 and 200 ms, respectively. All conditions were run in a counterbalanced order across subjects. There were 8 participants.

Results

As noted above, we report the difference between the adjusted duration of the comparison NBN and the veridical duration of the test NBN in each condition with the corresponding values obtained with no WBN subtracted out. Averaged across all 8 participants, these values were 36.0, 31.7, and 11.1 ms in conditions HHH, HHL, and HLH, respectively. A one-way ANOVA revealed a significant main effect of condition, $F(2, 14) = 6.7$, $p < .03$, $\eta^2 = .491$. Crucially, the adjusted duration was longer in condition HHL than in condition HLH (two-tailed *t* test, $df = 7$, $p < .01$), indicating that the main effect of the WBN was to decrease the perceived duration of the test NBN rather than to increase that of the comparison. Condition HHL did not differ significantly from HHH, indicating that the frequency relationship between the WBN and the comparison NBN had no effect ($df = 7$, $p = .73$).

The difference between adjusted and veridical WBN was, even in condition HHH, smaller than in Experiments 1–3. Inspection of the individual data revealed that 2 participants who had not taken part in Experiments 1–3 did not adjust the comparison NBN to be longer than the test NBN, even in condition HHH. These were 2 of only 3 participants from the 33 tested in our experiments who did not show such an effect when the WBN, test NBN, and comparison NBNs all had the same center frequency (P11 and P14, shown in Table 1 in bold italics). When their data were excluded, the same overall pattern of results was obtained, but with larger values overall: 52.5, 39.2, and 15.2 ms. There was once more a significant main effect of condition, $F(2, 10) = 14.1$, $p = .001$; and condition HHL again differed significantly from HLH ($df = 5$, $p < .05$) and not from HHH ($df = 5$, $p = .131$).

Experiment 5: Electrophysiology

Rationale and Stimuli

Experiment 5 reports a neural correlate of the duration illusion using the MMN paradigm. The MMN is a negative wave that is generated in response to a rare "deviant" stimulus presented in a sequence of more frequently presented "standards." It can be observed without requiring the participant to make any overt response to the sound, and it has a latency that is typically between 100 and 200 ms. Although it is thought to have multiple generators, its major source has been shown to be in auditory areas along the supratemporal plane (Alho, Woods, Algazi, Knight, & Näätänen, 1994; Giard, Perrin, Pernier, & Bouchet, 1990; Javitt, Steinschneider, Schroeder, Vaughan, & Arezzo, 1994; Kasai et al., 1999; Kropotov et al., 1995). Obtaining a neural correlate of the illusion imposes some constraints on the stages of processing at which it occurs, rules out explanations based on changes in participants' response criteria, and paves the way for future studies investigating the dependence of the illusion on attention. Furthermore, as argued below, by correlating measures of the MMN latency with behavioral measures of changes in perceived duration, one can distinguish between two broad accounts of how the illusion is represented in the brain at the stage of processing where the MMN is generated.

Experiment 5A measured the latency of the MMN to an NBN presented immediately after either an on- or an off-frequency WBN. The difference between these two latencies is then correlated with the difference in the perceived duration of the NBN, obtained from a duration-adjustment task using identical stimuli and the same participants in Experiment 5B. If the NBN were registered as starting earlier at the stage of processing where the MMN is generated, then we might expect the MMN to be earlier in the on- than in the off-frequency conditions and expect this difference to correlate across participants with the difference between the same two conditions in Experiment 5B. That is, participants who heard the NBN as much longer in the on- than in the off-frequency condition of Experiment 5B should show much earlier MMNs in the on- than in the off-frequency condition of Experiment 5A. Conversely, if at the stage of processing where the MMN is generated the brain requires a longer time to identify the onset of the NBN, the opposite result might be obtained: That is, the MMN latency would be longer in the on-frequency than in the off-frequency condition, and participants who heard the NBN as much longer in the on- than in the off-frequency condition of Experiment 5B should show much later MMNs in the on- than in the off-frequency condition of Experiment 5. Note that the method of correlating results across participants controls for effects—such as the generally greater salience of the NBN onset in the on- than in the off-frequency condition—that might produce overall differences in latency between the MMNs in the two conditions,

Experiment 5A consisted of four conditions. In the on-frequency condition (Figure 5a), each 901-s sequence contained a series of 707–1414-Hz WBNs. The duration of each WBN was 300 ms, and the stimulus onset asynchrony (SOA) between WBNs was 850 ms. Seventeen percent of these WBNs—the deviants—were followed immediately by a 200-ms NBN whose spectrum spanned the 930–1072-Hz range. The remaining WBNs were termed standards. At the end of this sequence, the participant heard another sequence consisting solely of 200 repetitions of the stimuli defined as deviants in the main block, again presented with an SOA of 850 ms. The MMN was calculated by subtracting the mean response to these last 200 repeat deviants from that to the deviants in the main block. This measure, termed the *identity MMN* (Pulvermüller, Shtyrov, Ilmoniemi, & Marslen-Wilson, 2006), has the advantage of measuring the difference between the responses to two identical stimuli, differing only in the context in which they were presented.

The off-frequency condition (Figure 5b) was identical to the on-frequency condition, except that the WBN now spanned 2000–4000 Hz, so that it would be expected to produce a smaller change in the perceived duration of the immediately following NBN.

The "WBN duration" condition (Figure 5c) controlled for the possibility that an MMN could be generated, not in response to the start of the NBN but simply to the increase in the total duration of each sound. It was identical to the on-frequency condition, except that the deviants consisted of a single WBN with a duration of 500 ms—equal to the total duration of each deviant in the on-frequency WBN condition.

The early NBN condition (Figure 5d) was included as a check for the sensitivity of the experiment to any overall decreases in latency in the on-frequency condition. Specifically, if no significant difference in latency between the on-frequency and off-frequency conditions were observed, this could, in theory, have

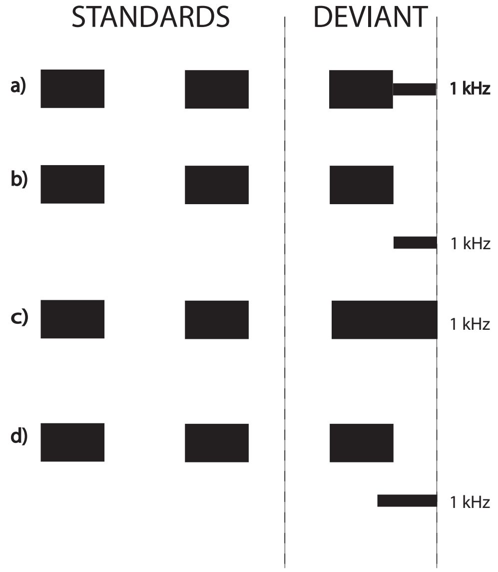

Figure 5. Schematic spectrograms of the stimuli used in Experiment 5.

been due to the experiment not being sensitive enough to detect a genuine difference in the time at which the NBN onsets were registered in those two conditions. The early NBN condition was identical to the off-frequency condition, except that the NBN started 80 ms earlier; by comparing the latencies of the MMNs in the off-frequency and early-NBN conditions, we could determine whether the experiment was sensitive enough to detect an 80-ms change in physical onset time. Note that in this condition the duration of the NBN was increased to 280 ms, so that the NBNs in all conditions ended at the same time relative to the WBN.

Each condition was run as a separate sequence, and the order of conditions was counterbalanced across the 16 participants. Following the EEG experiment, each participant completed a behavioral study (Experiment 5B) similar to Experiments 1–4. Experiment 5B consisted of three conditions, in each of which the test and comparison NBNs contained frequencies between 932 and 1072 Hz. In the on- and off-frequency conditions the bandwidth of the WBNs spanned 707 to 1414 Hz and 2000 to 4000 Hz, respectively, as in the corresponding conditions of Experiment 5A. In the no-WBN condition the WBN was absent. In both parts of Experiment 5, the stimuli were played out at a sampling rate of 44100 Hz using a Soundblaster PC card and presented to the left earpiece of a Sennheiser HD250 headset. The overall levels were 52, 59, and 54 dB SPL for the NBN, on-frequency WBN, and off-frequency WBN, respectively. The corresponding spectrum levels were 30, 30, and 21 dB SPL.

Electrophysiological Recordings

Subjects were comfortably seated in an electrically shielded and acoustically insulated chamber. They were instructed to listen passively and were not required to make any response. The EEG was recorded with Ag/AgCl electrodes mounted in an extended 10-20 system cap (EasyCap, Herrsching-Breitbrunn, Germany) with a 64-channel EEG setup (Neuroscan, Charlotte, NC). Cz was used as the reference electrode for recordings, with electrode AFz as the ground. Horizontal eye movements were monitored using electrodes (the positions of which were near the outer canthi of left and right eyes) on the cap. Vertical eye movements were monitored using two extra electrodes placed below and above the left eye. Signals were amplified, sampled at a rate of 500 Hz, and bandpass filtered between 0.1 and 100 Hz.

The acquired EEG traces were stored on the hard disk of a computer, and further processing was carried out off-line in the digital domain, starting with bandpass filtering (2–20 Hz, 24 dB/octave slopes). Event-related potentials were obtained by averaging epochs, which started 100 ms before and ended 850 ms after the onset of each WBN (in other words, the analysis period was 950 ms, including a 100-ms prestimulus baseline and extending 350 ms after the offset of the last stimulus). Epochs containing voltage variations in excess of 100 μV at any EEG or EOG channel were discarded. All epochs corresponding to a standard stimulus that just followed a deviant stimulus were excluded.

After the average across channels was subtracted from all traces (average referencing), the traces were re-referenced to the average of channels TP9 and TP10, which were close to the left and right mastoids. Finally, the data were re-referenced to the 100-ms baseline preceding the WBNs.

Results

In Experiment 5B, the adjusted duration of the comparison noise, minus the veridical duration of the test NBN, was 78.1 ms in the on-frequency WBN condition, 15.1 ms in the off-frequency WBN condition, and 2.8 ms in the no-WBN condition. This confirms the findings of Experiments 1–4 that the on-frequency WBN increases the perceived duration of the test NBN and is significantly more effective than an off-frequency WBN at doing so.

The identity MMNs from the on-frequency WBN of Experiment 5A are shown, for all electrodes, in Figure 6A. It can be seen that the maximal negative deflections occur over fronto-central electrodes and that, as is typical of the MMN, the polarity inverts below the mastoids. This pattern was observed in all conditions. In the remainder of this article we consider the MMN as measured on electrode FCz, as shown for each condition in Figure 6B. The time values on the abscissa are expressed relative to the onset of the WBN, and the arrows indicate the onset time of the NBN; this also corresponds to the time at which the deviants differ from the standards. Note that in the on-frequency, off-frequency, and early NBN conditions, the MMN consists of two peaks separated by approximately 180 ms. Such double-peaked MMNs have been observed in the past (Näätänen, Pakarinen, Rinne, & Takegata, 2004; Pulvermüller & Shtyrov, 2003; Shtyrov & Pulvermüller, 2007), although the reasons for the presence of more than one peak are not known. The latencies and amplitudes of these two peaks, as well as those of the intervening positive peak, are shown in Table 2.

A comparison of the black and green curves in Figure 6B show that the MMN peaks occur slightly later in the on-frequency than in the off-frequency condition. To measure the latency for each participant and condition, we adopted a procedure that evaluated peaks over the same time window for both conditions and made minimal assumptions about the timing of each peak. Specifically, we searched for the two most negative values over the time window 350–750 ms, with the constraint that the two peaks were separated by at least 100 ms.[2] The latencies of Peaks 1 and 2 were 455 and 636 ms for the on-frequency WBN condition and 415 and 599 ms for the off-frequency WBN condition. Paired-sample t tests revealed that the latencies of both Peaks 1 and 2 were significantly longer in the on-frequency than in the off-frequency condition ($df = 15$, $p < .01$ and $p < .001$, respectively).

It is possible that the difference in mean latency between the on- and off-frequency conditions is due to the onset of the NBN being more salient in the latter condition and that this occurs for reasons unrelated to the duration illusion. However, more importantly, the difference in the latency of the first MMN peak between the on- and off-frequency conditions of Experiment 5B correlated positively across participants with the difference in perceived duration between the two analogous conditions of

---

[2] Note that the mean values of these peaks, obtained separately for each participant and then averaged, will not correspond exactly to the values shown in Table 2, which were obtained by averaging the traces for each participant and then measuring the peaks by inspection of the resulting average curves.

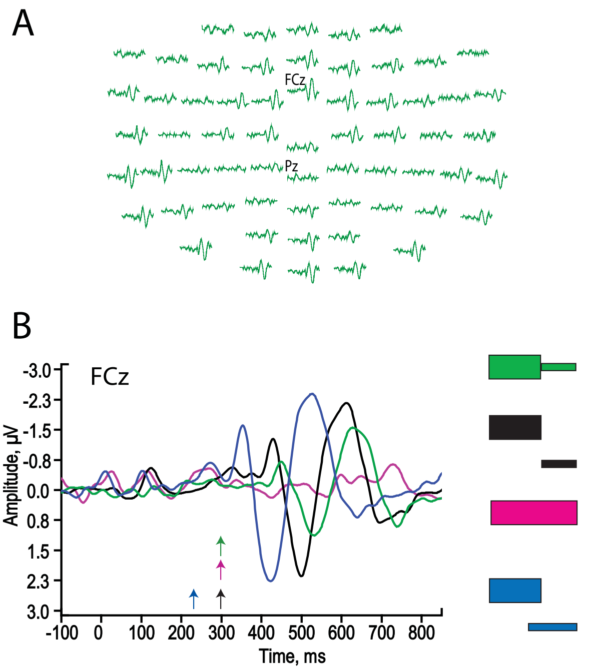

Figure 6. Panel A shows the identity mismatch negativity (MMN) traces obtained on all electrodes, averaged across participants, from the on-frequency wideband noise condition of Experiment 5. Panel B shows the identity MMNs recorded from electrode FCz in each of the four conditions, with a schematic reminder of the deviant stimulus in each condition shown in a matching color on the right.

Experiment 5A ($r = .53$, $df = 15$, $p < .05$; Figure 7).[3] Hence listeners who perceived the duration of the NBN to be longer in the on-frequency than in the off-frequency condition of Experiment 5B tended to show a *later* first MMN peak in that condition of Experiment 5A. This across-participant correlation demonstrates that, although the direction of causality remains to be determined, some common process affects both the duration illusion and the latency of the MMN. This processing does not appear to reflect a bottom-up effect, whereby the response to the NBN occurs earlier under conditions where the illusion occurs, because the direction of the correlation is such that a larger duration illusion is accompanied by later MMNs in the on-frequency, compared to the off-frequency, condition. Rather, the results are more consistent with the processing required to register the NBN taking longer in the on-frequency WBN condition, where the perceived NBN duration is increased by the WBN, than in the off-frequency WBN condition, where it is not.

When the NBN is turned on 80 ms earlier (the early NBN condition; blue curves), Peaks 1 and 2 in the averaged curves occur at 352 and 526 ms, respectively, which are 96 and 88 ms earlier than the corresponding peaks in the off-frequency WBN condition (Table 2). The fact that both peaks in the early NBN condition were about 80 ms earlier than in the off-frequency WBN condition shows that they are both likely to be initiated by the onset of the NBN, rather than Peak 2 being elicited by the NBN offset (which occurred at the same time in the two conditions).[4]

Experiment 6: Studying Perceived Onset Time via Rhythm Judgments

Rationale and Overview

Experiment 6 examined whether the changes in perceived duration observed in Experiments 1–4 were accompanied by corresponding changes in onset time, as estimated at the stage of processing where rhythmic judgments are made. It built on previous evidence that, when judging whether a sequence of sounds is isochronous, listeners base their judgments on the onsets, rather than the offsets, of each sound (Vos, 1977; Vos, Mates, & Kruysbergen, 1995). A schematic of the baseline (“no WBN”) condition is shown in Figure 8a. Each trial consisted of six 50-ms 1000-Hz pure tones separated from each other by the same interstimulus interval (ISI), which could be 500, 550, or 600 ms. The final stimulus in the train was a 200-ms 2828-Hz pure tone. The participant’s task was to adjust the onset time of this test tone so that it was isochronous with the preceding tones (all of which had the same ISI). In the on-frequency WBN condition (Figure 8b), the test tone was immediately preceded by a 300-ms, 2000–4000 Hz WBN. Participants could adjust the onset time of the WBN and test tone together so that the test tone sounded isochronous with the preceding tones, and they were instructed to ignore the noise. The rationale was that if the on-frequency WBN caused the test tone to be heard as starting earlier, then, to compensate for this, participants should adjust it to start later than when no WBN was present. They should also adjust it to start later than in the off-frequency WBN condition (Figure 8c), in which the WBN spanned frequencies from 707–1414 Hz. Nine normal-hearing listeners took part.

Experiment 6B was similar in design to Experiments 1–4 but used stimuli similar to those of Experiment 6A. In the on-frequency WBN condition, a 200-ms 2848-Hz test tone was immediately preceded by a 300-ms, 2000–4000 Hz WBN. The participant’s task was to adjust the duration of a 2848-Hz comparison tone to match the perceived duration of the test. In the off-frequency WBN and no WBN conditions, the WBN spanned 707–1414 Hz or was absent. Twelve normal-hearing listeners took part, 8 of whom had also taken part in Experiment 6A. Experiment 6A was performed before Experiment 6B, for all participants.

Method and Procedure

In both parts of the experiment, the stimuli were played out of an SB Live! Sound card at a sampling rate of 44 100 Hz, low-pass filtered at 20000 Hz (Kemo VBF/01), attenuated (TDT PA4), and presented via a headphone amplifier to one earpiece of a

---

[3] The correlation with Peak 2 was not significant ($r = -.16$, $df = 15$, *ns*).

[4] Two further observations are worthy of note. First, the amplitude of the MMN was generally similar across the on-frequency, off-frequency, and early-NBN conditions. An exception is the fact that Peak 1 was significantly smaller in the on-frequency WBN condition than in the off-frequency WBN condition (paired-sample t test, $df = 15$, $p < .05$). Second, in the WBN-duration condition, the only deviation observed in the difference curve was a small negative deflection with a latency of about 720 ms (420 ms after the WBN ended for the standard stimuli). One reason for the long latency of this response may be that it was elicited by the end of the deviant WBN, which occurred at 500 ms, rather than by the deviant not ending at 300 ms (as occurred for the standard stimuli).

Table 2

Latencies and Amplitudes of the Major Peaks Shown in Figure 6B, Based on Data Averaged Across All Participants

Condition	Negative Peak 1		Negative Peak 2		Positive peak	
	Latency (ms)	Amplitude (μV)	Latency (ms)	Amplitude (μV)	Latency (ms)	Amplitude (μV)
On-frequency WBN	448	−0.699	628	−1.551	532	1.140
Off-frequency WBN	428	−1.276	612	−2.167	498	2.155
WBN duration	—	—	—	—	—	—
Early NBN	352	−1.607	526	−2.396	424	2.278

Note. The latencies and amplitudes shown do not necessarily correspond to those obtained when the peak values are selected separately for each participant and then averaged. The dashes shown for the wideband noise (WBN) duration condition reflect the fact that no clearly identifiable peaks were observed. NBN = narrowband noise.

Sennheiser HD250 headset. All tones and WBNs were turned on and off with 5-ms raised-cosine ramps. The overall levels of the on- and off-frequency WBNs were 61 and 65 dB SPL, respectively, and the level of each tone was 51 dB SPL. The level of the tone was 23 dB higher than the spectrum level of the on-frequency WBN. This value is the same as the difference between the overall level of the NBN and the spectrum level of the on-frequency WBN in Experiments 1–3.

Each run of Experiment 6A consisted of a sequence of presentations of a single condition and ISI (500, 550, or 600 ms). The first presentation consisted of the five 50-ms tones presented at the appropriate ISI, followed by the test tone. The delay between the start of the last 50-ms tone and the start of the test tone was drawn at random from a rectangular distribution spanning the range 0.5 to 1.5 times the nominal ISI. At the end of each presentation the participant could increase or decrease the value of this delay by 10 or 40 ms by clicking on one of four virtual buttons on a computer screen. This delay was then used for the next presentation. When satisfied that all tones sounded isochronous, the participant clicked on a fifth button and the final adjusted value was recorded. A total of at least 16 such runs were obtained for each combination of condition and ISI, and the results from all runs were averaged. Runs were performed in a counterbalanced order.

The procedure for Experiment 6B was similar to that of Experiments 1–4. After each presentation of the WBN, test, and comparison sounds, the participant could adjust the duration of the comparison sound to be presented on the subsequent trial. Unlike Experiments 1–4, and similar to Experiment 6A, adjustments were made by clicking on virtual buttons on a computer screen. The method of adjustment followed the same procedure as described in Experiment 4, with order of presentation counterbalanced across subjects and with the mean of between 12 and 25 judgments (depending on the time availability of each participant) used to represent each condition for each participant.

Results

The results of Experiment 6B showed that the adjusted duration of the comparison tone minus the veridical duration of the test tone was 73.3, 16.8, and 15.6 ms in the on-frequency, off-frequency, and no-WBN conditions, respectively. Hence, the difference in the adjusted duration between the on-frequency and off-frequency WBN conditions was 56.5 ms. This effect was therefore of a similar size to that observed in Experiments 1–4 with an NBN instead of a tone. For the 8 participants who also took part in Experiment 6A, the mean difference between the on- and off-frequency conditions was 55.7 ms.

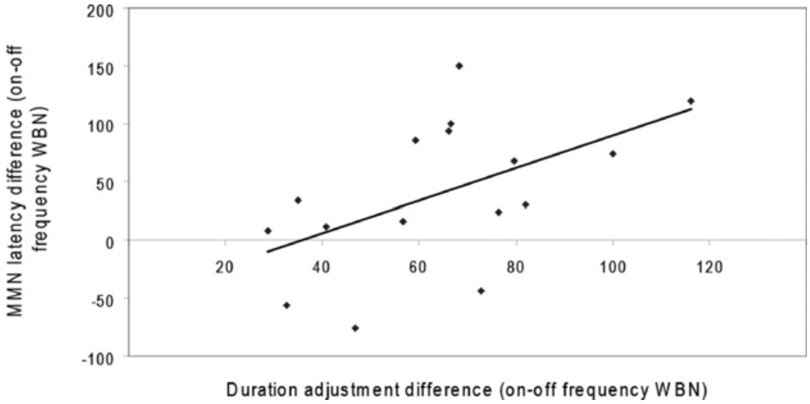

Figure 7. The abscissa shows the difference in adjusted duration between the on- and off-frequency conditions for each participant in Experiment 5B. The ordinate shows the difference in mismatch negativity (MMN) latency between the on- and off-frequency condition for each participant in Experiment 5A. Each point shows the data for one participant. The solid line shows the best-fitting linear regression to the data. WBN = wideband noise.

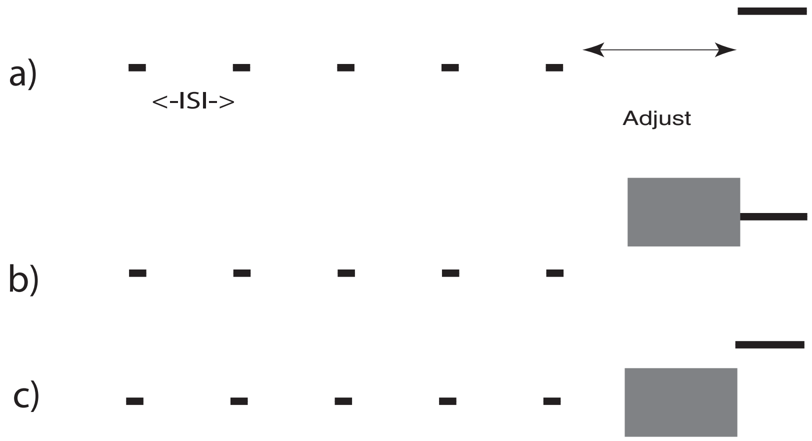

Figure 8. Schematic spectrogram of the stimuli used in the silence, on-frequency, and off-frequency conditions of Experiment 6A. ISI = interstimulus interval.

It is also worth noting that 2 participants in this experiment failed to show a difference between the two conditions. These two participants were P11, who also failed to show the duration illusion in Experiment 4, and P26. As noted in the description of the results of Experiment 4, only 3 participants out of a total of 31 tested in our experiments failed to show the duration illusion, and the fact that P11 failed to show it in two different experiments, which differed in the narrowband stimulus used (NBN vs. tone), is further evidence for the stability of the effect and of individual differences in its potency.

Each condition of Experiment 6A is represented in Figure 9 by a cluster of three bars, where each bar in the cluster represents data obtained at one ISI. The values shown represent the adjusted ISI between the last 50-ms tone and the test tone, minus the nominal ISI, averaged across participants. In the no-WBN condition these values are all negative, suggesting that there is a response bias toward adjusting the test tone slightly too early. A similar bias is apparent in the data obtained with the off-frequency WBN. In contrast, the results obtained with an on-frequency WBN all correspond to positive values. Hence participants adjusted the onset of the test tone to be later than in the other conditions, consistent with the tone being judged to start earlier. The results of Experiment 6A therefore suggest that at the stage of processing where listeners judge rhythm, the tone is judged as starting earlier under conditions where it is perceived as being longer. Averaged across participants and ISIs, the difference between the adjusted onset times in the on- and off-frequency conditions was 34.9 ms. For the 8 participants in common with Experiment 6B, this value was 31.8 ms.

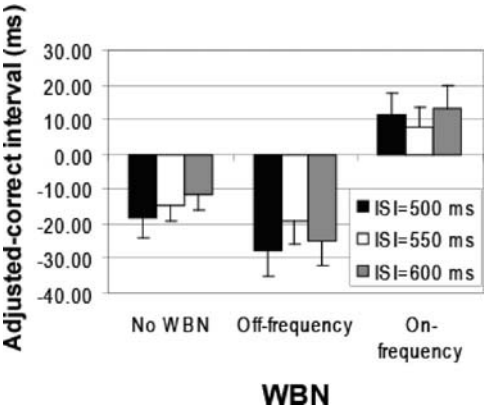

Figure 9. Difference between the adjusted and isochronous interval averaged across all participants in Experiment 6. Each cluster of three bars shows the data obtained in one condition at ISIs of 500, 550, and 600 ms, respectively. The error bars show standard errors calculated with sources of across-listener variation removed by subtracting the average adjustment across conditions for a given participant from each data point for that participant. WBN = wideband noise.

Interim Discussion

The rhythm judgments obtained in Experiment 6 could potentially have been influenced by two biases, which we consider briefly here.

The three ISIs used in Experiment 6A were chosen to produce SOAs between the 50-ms tones of 550, 600, and 650 ms. The values were selected to be close to 600 ms, as this corresponds to the SOA at which rhythm judgments are generally most accurate (Fraisse, 1984; Parncutt, 1994). Because the WBN had a duration of 300 ms, its onset would have bisected an SOA of 600 ms between the final 50-ms tone and the 200-ms test tone, possibly influencing the results. However, it can be seen from Figure 9 that the pattern of results is similar for all SOAs, suggesting that this "bisection" did not substantially influence the results reported here.

In the baseline and off-frequency conditions of Experiment 6A, listeners adjusted the target tone to start slightly earlier than the physically isochronous SOA. This is unlikely to be due to the target tone having a different frequency from the preceding tones, because a similar bias occurred in a pilot experiment where all tones had the same frequency. One possible explanation is suggested by the results of a recent study by Schimmel and Kohlrausch (2008), who required participants to adjust the position of a target noise burst, presented in the middle of a sequence of marker bursts, so that all sounds were perceived as isochronous. When all bursts had the same duration the resulting adjustments corresponded to physical isochrony, but when the target and marker durations were 350 ms and 50 ms, respectively, the target was adjusted to start 22 ms earlier than the physically isochronous point. It may well be, then, that our 200-ms target was adjusted to start earlier than physical isochrony because its duration was longer than that of the 50-ms markers. Note, however, that this cannot account for the main finding that the tone was adjusted as starting later when preceded by an on-frequency WBN, because that WBN increased the perceived duration of the tone.

Discussion

Basis of the Duration Illusion

The experiments reported here describe an increase in the perceived duration of a narrowband stimulus produced by a preceding WBN. An estimate of the size of this duration illusion can be obtained by inspecting the results of Experiments 5B and 6B, which included the largest numbers of participants; these values were 78 and 73 ms, respectively, which reduce to 63 and 56 ms if one subtracts the size of the effect obtained with an off-frequency WBN. Taken together, although the results of the experiments do not provide a complete characterization of the neural mechanisms responsible, they do collectively impose some constraints on the processes underlying the illusion.

First, Experiment 4 showed that the illusion reflects a genuine increase in the perceived duration of the target NBN, rather than being due to some "long range" effect on the subsequent comparison NBN, as has been previously reported in the loudness domain (Scharf et al., 2002). Second, it does not appear to be true that the WBN causes the latency of early neural responses to the NBN to be shorter under conditions where the illusion occurs. Rather, Experiment 5 showed that participants who show the largest difference in perceived duration between on- and off-frequency conditions tend to show later MMNs in the on-frequency condition. Finally, the fact that correlates of the illusion can be obtained using electrophysiological recordings (Experiment 5) and perceptual rhythm judgments (Experiment 6) shows that the illusion does not require subjects to make an explicit judgment of duration, and is likely to reflect a genuine change in perception rather than being due to a response bias.

The results described here also lay the groundwork for a more positive hypothesis of the mechanisms underlying the illusion. The fact that the MMN latency is not only longer in the on-frequency WBN condition than in the off-frequency WBN condition of Experiment 5A, but also that this difference correlates across subjects with the difference in perceived duration, suggests that some processes involved in the illusion have occurred by the time (approximately 130–150 ms) and place (auditory cortex) corresponding to the generation of the MMN (Alho et al., 1994; Giard, Perrin, Pernier, & Bouchet, 1990; Javitt et al., 1994; Kasai et al., 1999; Kropotov et al., 1995). This correlation is consistent with some additional computation being required to register the presence of the NBN when preceded by an on-frequency WBN than by an off-frequency WBN. One reason for this may be that whereas in the off-frequency WBN condition it is sufficient to detect energy occurring in a new frequency region (Figure 8c), in the on-frequency WBN condition it is necessary to compare the energy in the region corresponding to the WBN to that encompassed by the NBN (Figure 8b). Whatever the nature of that additional computation, other aspects of auditory processing appear capable of interpreting the NBN as starting earlier when preceded by an on-frequency WBN, as evidenced by the rhythmic judgments of Experiment 6. In summary, the answer to the question "Is the NBN encoded as starting earlier under conditions where the illusion causes it to be heard as longer?" may well depend on the stage of encoding one is considering. Our results provide no evidence that an earlier neural response occurs from the bottom up, but show that, at least when rhythmic judgments are made, the perceived duration increase does appear to be accompanied by the NBN being heard as starting earlier.

Implications for Models of Duration Encoding

Many models of duration encoding (e.g., Creelman, 1962) assume that the brain counts the number of internally generated "pulses" between the start and end of a stimulus (for a review, see Allan, 1979).What these models share is the assumption that initiation and termination of an internal counting process is determined by markers derived from the start and end of the physical stimulus. Recently, Tsuzaki and colleagues (Tsuzaki & Kato, 2000; Tsuzaki & Tanaka, 2003) have proposed a modification to this general class of model. In one experiment, Tsuzaki and Kato (2000) required participants to judge the total duration of a target tone that could be uninterrupted, or could contain a silent gap that was optionally filled by a replacing sound. They reported that total perceived duration of the target was smallest in the condition with the silent gap, increased progressively as replacing sounds that induced greater degrees of illusory continuity were introduced, and longest for the uninterrupted tone. They concluded that the gating of the internal counting process is controlled, in a continuous manner, according the degree of perceptual evidence for the target sound. Our findings could be incorporated into this model if it is assumed that the gate is partially open during the portion of the WBN preceding the NBN onset. Because no evidence for the NBN occurs until the WBN has stopped, one would further have to assume that the gate was not controlled in real time and that the calculation of duration occurred retrospectively.

Relationship to the Continuity Illusion

As noted in the introduction, the duration illusion described here has several aspects in common with the continuity illusion. At a general level, both phenomena reflect a change in the perceived temporal characteristics of a stimulus when its temporal boundaries are made ambiguous by the presence of an inducing sound. Indeed, as illustrated in Figure 1A, the stimuli used to measure the duration illusion resemble those used to study the continuity illusion, except that in the former case, the "inducee" (i.e., the NBN) occurs only after the inducer. Furthermore, both phenomena depend in a broadly similar way on the presence or absence of temporal gaps between the inducer and inducee and on the frequency relationship between them. Finally, correlates of both the continuity illusion (Micheyl et al., 2003) and the duration illusion can be observed in the MMN.

The difference between the two phenomena is, as noted above, that the continuity illusion is usually studied using stimuli in which the inducee is presented both before and after the inducer. If, for example, the continuity illusion arose from neural activity, similar to that produced by the inducee alone, persisting through the inducer then this would be evidence that the continuity and duration illusions arose from different underlying mechanisms. However, there is convincing evidence that this is not the case. For example, Bregman and Dannenbring (1977) showed that the continuity illusion could be reduced by momentarily changing the amplitude of the inducee immediately before and after the inducer was turned on and off. The fact that this reduction could occur even when the amplitude of the inducee was momentarily *increased* argues strongly against the persistence hypothesis, and is more consistent with an account based on the ambiguity of the temporal boundaries of the inducee. We therefore conclude that the most parsimonious and convincing explanation is that the duration and continuity illusions are related, and both reflect the way in which the auditory system deals with situations where the onsets and offsets of sounds are ambiguous.

Summary

A series of six experiments studied a phenomenon in which the perceived duration of an NBN can be increased by an immediately preceding WBN of a similar frequency. This duration illusion increases as the duration of the NBN is lengthened from 50 to about 500 ms and as the duration of the WBN is lengthened from 100 ms to about 300 ms. The maximal increase in perceived duration is about 60–80 ms. The effect disappears, on average, when the gap between the WBN and NBN is as long as 20–40 ms, but can persist at longer gaps (80 ms) for some participants. A similar effect occurs when the NBN is replaced by a tone. Artefactual effects, akin to those that may affect experiments in the loudness domain, were ruled out.

The perceived duration increase was associated with an increase in the latency of the MMN to the NBN, and participants who heard the NBN as much longer when preceded by an on-frequency than by an off-frequency WBN tended to show a longer MMN latency in the on- than in the off-frequency condition. Hence at this stage of processing we could find no evidence that the onset of the NBN was encoded as starting earlier under conditions where it was judged as longer. Rather, those results are consistent with some additional processing being required to register the existence of an NBN after an on-frequency WBN had been turned off. In contrast, the results of an experiment involving judgments of rhythm indicate that at this stage of processing, the NBN is heard as starting earlier under conditions where it is judged as being longer.

References

Alho, K., Woods, D. L., Algazi, A., Knight, R. T., & Näätänen, R. (1994). Lesions of frontal-cortex diminish the auditory mismatch negativity. *Electroencephalography and Clinical Neurophysiology, 91,* 353–362.

Allan, L. G. (1979). The perception of time. *Perception and Psychophysics, 26,* 340–354.

Brand, A., Urban, A., & Grothe, B. (2000). Duration tuning in the mouse auditory midbrain. *Journal of Neurophysiology, 84,* 1790–1799.

Bregman, A. S. (1990). *Auditory scene analysis.* Cambridge, MA: MIT Press.

Bregman, A. S., & Dannenbring, G. L. (1977). Auditory continuity and amplitude edges. *Canadian Journal of Psychology, 31,* 151–159.

Broadbent, D. E., & Ladefoged, P. (1959). Auditory perception of temporal order. *Journal of the Acoustical Society of America, 31,* 1539.

Carlyon, R. P. (1988). The development and decline of forward masking. *Hearing Research, 32,* 65–80.

Carlyon, R. P., Micheyl, C., Deeks, J. M., & Moore, B. C. J. (2004). Auditory processing of real and illusory changes in FM phase. *Journal of the Acoustical Society of America, 116,* 3629–3639.

Ciocca, V., & Bregman, A. S. (1987). Perceived continuity of gliding and steady-state tones through interrupting noise. *Perception and Psychophysics, 42,* 476–484.

Creelman, C. D. (1962). Human discrimination of auditory duration. *Journal of the Acoustical Society of America, 34,* 582–593.

Fraisse, P. (1984). Perception and estimation of time. *Annual Review of Psychology, 35,* 1–36.

Giard, M. H., Perrin, F., Pernier, J., & Bouchet, P. (1990). Brain generators implicated in the processing of auditory stimulus deviance: A topographic event-related potential study. *Psychophysiology, 27,* 627–640.

Glasberg, B. R., & Moore, B. C. J. (1990). Derivation of auditory filter shapes from notched-noise data. *Hearing Research, 47,* 103–138.

He, J., Hashikawa, T., Ojima, H., & Kinouchi, Y. (1997). Temporal integration and duration tuning in the dorsal zone of cat auditory cortex. *Journal of Neuroscience Methods, 17,* 2615–2625.

Houtgast, T. (1972). Psychophysical evidence for lateral inhibition in hearing. *Journal of the Acoustical Society of America, 51,* 1885–1894.

Javitt, D. C., Steinschneider, M., Schroeder, C. E., Vaughan, H. G., & Arezzo, J. C. (1994). Detection of stimulus deviance within primate primary auditory-cortex: Intracortical mechanisms of mismatch negativity (MMN) generation. *Brain Research, 667,* 192–200.

Kasai, K., Nakagome, K., Itoh, K., Koshida, I., Hata, A., Iwanami, A., et al. (1999). Multiple generators in the auditory automatic discrimination process in humans. *NeuroReport, 10,* 2267–2271.

Kropotov, J. D., Naatanen, R., Sevostianov, A. V., Alho, K., Reinikainen, K., & Kropotova, O. V. (1995). Mismatch negativity to auditory stimulus change recorded directly from the human temporal cortex. *Psychophysiology, 32,* 418–422.

Lyzenga, J., Carlyon, R. P., & Moore, B. C. J. (2005). Dynamic aspects of the continuity illusion: Perception of level, and of the depth, rate, and phase of modulation. *Hearing Research, 210,* 30–41.

Micheyl, C., Carlyon, R. P., Pulvermüller, F., Shtyrov, Y., Hauk, O., & Dodson, T. (2003). Neurophysiological correlates of a perceptual illusion: A mismatch negativity study. *Journal of Cognitive Neuroscience, 15,* 747–758.

Näätänen, R., Pakarinen, S., Rinne, T., & Takegata, R. (2004). The mismatch negativity (MMN): Towards the optimal paradigm. *Clinical Neurophysiology, 115,* 140–144.

Parncutt, R. (1994). A perceptual model of pulse salience and metrical accent in musical rhythms. *Music Perception, 11,* 409–464.

Pulvermüller, F., & Shtyrov, Y. (2003). Automatic processing of grammar in the human brain as revealed by the mismatch negativity. *NeuroImage, 20,* 159–172.

Pulvermüller, F., Shtyrov, Y., Ilmoniemi, R. J., & Marslen-Wilson, W. D. (2006). Tracking speech comprehension in space and time. *NeuroImage, 31,* 1297–1305.

Scharf, B., Buus, S., & Nieder, B. (2002). Loudness enhancement: Induced loudness reduction in disguise? *Journal of the Acoustical Society of America, 112,* 807–810.

Schimmel, O., & Kohlrausch, A. (2008). On the influence of interaural differences on temporal perception of noise bursts of different durations. *Journal of the Acoustical Society of America, 123,* 986–997.

Shtyrov, Y., & Pulvermüller, F. (2007). Early MEG activation dynamics in the left temporal and inferior frontal cortex reflect semantic context integration. *Journal of Cognitive Neuroscience, 19,* 1633–1642.

Tsuzaki, M., & Kato, H. (2000). Shrinkage of perceived tonal duration produced by extra sounds: Effects of spectral density, temporal position, and transition direction. *Perception, 29,* 989–1004.

Tsuzaki, M., & Tanaka, M. (2003). Shrinkage in the perceived duration of speech and tone by acoustic replacement. *Japanese Psychological Research, 45,* 129–139.

Vicario, G. (1960). L'effetto tunnel acustico [The acoustic tunnel effect]. *Rivista di Psicologia, 54,* 41–52.

Vliegen, J., Moore, B. C. J., & Oxenham, A. J. (1999). The role of spectral and periodicity cues in auditory stream segregation, measured using a temporal discrimination task. *Journal of the Acoustical Society of America, 106,* 938–945.

Vos, P. G. (1977). Temporal duration factors in the perception of auditory rhythmic patterns. *Scientific Aesthetics, 1,* 183–199.

Vos, P. G., Mates, J., & van Kruysbergen, N. W. (1995). The perceptual centre of a stimulus as the cue for synchronization to a metronome: Evidence from asynchronies. *Quarterly Journal of Experimental Psychology, 48,* 1024–1040.

Warren, R. M. (1999). *Auditory perception: A new analysis and synthesis.* Cambridge, UK: Cambridge University Press.

Warren, R. M., Obuseck, C. J., & Ackroff, J. M. (1972). Auditory induction: Perceptual synthesis of absent sounds. *Perception and Psychophysics, 20,* 380–386.

Xia, Y.-F., Qi, Z.-H., & Shen, J.-X. (2000). Neural representation of sound duration in the inferior colliculus of the mouse. *Acta Otolaryngologica, 120,* 638–643.

Zwislocki, J. J., Piroda, E., & Rubin, H. (1959). On some poststimulatory effects at the threshold of audibility. *Journal of the Acoustical Society of America, 31,* 9–14.

Received March 26, 2008
Revision received November 1, 2008
Accepted November 23, 2008 ■