Published in final edited form as:

J Exp Psychol Hum Percept Perform. 2014 June ; 40(3): 908–914. doi:10.1037/a0035411.

Perceptual asymmetry induced by the auditory continuity illusion

Dorea R. Ruggles and Andrew J. Oxenham

University of Minnesota, Minneapolis, MN

Abstract

The challenges of daily communication require listeners to integrate both independent and complementary auditory information to form holistic auditory scenes. As part of this process listeners are thought to “fill in” missing information to create continuous perceptual streams, even when parts of messages are masked or obscured. One example of this filling-in process – the auditory continuity illusion – has been studied primarily using stimuli presented in isolation, leaving it unclear whether the illusion occurs in more complex situations with higher perceptual and attentional demands. In this study, young normal-hearing participants listened for long target tones, either real or illusory, in “clouds” of shorter masking tone and noise bursts with pseudo-random spectro-temporal locations. Patterns of detection suggest that illusory targets are salient within mixtures, although they do not produce the same level of performance as the real targets. The results suggest that the continuity illusion occurs in the presence of competing sounds, and can be used to aid in the detection of partially obscured objects within complex auditory scenes.

Keywords

continuity illusion; auditory object; perceptual search; perceptual asymmetry

Introduction

The continuity illusion occurs when a masked, or obscured, portion of a stimulus is perceptually “filled in” to create the illusion of a continuous stream of information (Bregman, 1990; Warren, 1999). Conditions that foster this type of filling in have been identified in tactile (Kitagawa, Igarashi, & Kashino, 2009), visual (Komatsu, 2006), and auditory perception (King, 2007). In audition, the induction of missing information can play a role in speech understanding (Bashford, Riener, & Warren, 1992; Shahin, Bishop, & Miller, 2009; Shinn-Cunningham & Wang, 2008), and has been studied because of its potential for providing information about the perceptual and neurological aspects of auditory object formation. Early studies by Hougast (1972) and Duifhuis (1980) used the continuity illusion in the form of pulsation thresholds to demonstrate psychophysical correlates of nonlinear frequency tuning in the auditory periphery, and later studies have examined the neural correlates of the continuity illusion at higher levels of the auditory system using

neuroimaging (Riecke, Vanbussel, et al., 2012; Riecke, van Opstal, Goebel, & Formisano, 2007; Shahin et al., 2009).

The conditions under which the continuity illusion occurs have been studied since it was initially identified (Miller & Licklider, 1950). It is generally believed that the illusion occurs when the peripheral auditory response (e.g., auditory-nerve activity) produced by the interfering sound (or masker) overlaps completely with the response produced by the target sound (e.g., Warren, Obusek, & Ackroff, 1972; Duifhuis, 1980; Houtgast, 1972; Petkov & Sutter, 2011). Our understanding of the conditions necessary for the illusion have been refined by recent studies, which have shown that the illusion can still occur under some circumstances in which the peripheral auditory response provides evidence of the interruption, suggesting that masking of the interruption's onset and offset are more critical than the ongoing portion (Haywood, Chang, & Ciocca, 2011) or that global features, such as the specific loudness of the interferer, play a more dominant role than the interferer's fine-grained temporal structure (Riecke, Micheyl, & Oxenham, 2012).

The physiological basis of the continuity illusion has also been studied using a wide range of electrophysiological techniques that have revealed important details of its generation and attentional requirements. Especially significant for the current study is the finding by Micheyl et al. (2003), who used mismatched negativity (MMN) methods to show that physiological responses consistent with the continuity illusion do not seem to depend on focused attention. A similar finding was presented by Heinrich et al. (2011), who found that fMRI correlates of the continuity illusion seem to be independent of attention for complex vowel-like stimuli.

These studies provide some neurophysiological evidence that the continuity illusion is represented neurally for both simple and complex sounds in a way that may not depend on directed attention. Such findings would be strengthened via behavioral evidence that the continuity illusion effectively generates relevant auditory objects within attentionally demanding and/or complex acoustic environments (Gutschalk, Micheyl, & Oxenham, 2008; Jones, Macken, & Murray, 1993).

To investigate the role of the continuity illusion in auditory mixtures we used an auditory perceptual asymmetry identified by Cusack and Carlyon (2003), analogous to findings in the visual modality (e.g., Treisman & Gelade, 1980). Cusack and Carlyon (2003) found that long tones in mixtures of short tones were detected more easily than short tones in mixtures of long tones, and attributed these asymmetries to the existence of feature-specific neurons tuned to longer, rather than shorter, durations.

We asked whether an illusory long tone, comprised of two short tones interrupted by a noise burst, would be detected if it were embedded in a complex pattern of similar but non-contiguous short tones and noise bursts. The question of whether illusory long tones evoke the same feature mapping and detection asymmetries as actual long tones has the potential to contribute to a deeper understanding of the continuity illusion and the processes of feature coding and selection in complex acoustic environments. If the results show that the illusion is not detectable in complex mixtures of tones and noises and produces no perceptual

asymmetries, we may conclude that the continuity illusion, as measured behaviorally, stems from processes that are secondary to the feature mapping that results in auditory asymmetry and is thus unlikely to play an important role in object formation in complex acoustic environments. In contrast, if listeners are able to detect illusory long tones in mixtures of tones and noise and display a perceptual asymmetry similar to that found for physical long tones, it would suggest that the continuity illusion is formed prior to, or in conjunction with, the feature mapping associated with perceptual asymmetries, and so could play a crucial role in parsing complex auditory scenes.

Experiment 1

Methods

The experiment tested listeners' ability to detect illusory long tones elicited by a continuity illusion when the target tones were embedded in "clouds" of distracting tones and noises. Five conditions were studied (see Figure 1). In all conditions, "short" and "long" tones had total durations of 100 ms and 300 ms, respectively. All the noise bursts had total durations of 100 ms. Raised-cosine onset and offset ramps were applied to the first and last 10 ms of the tone and noise bursts. Pure-tone frequencies were randomly selected from 1/3 octave ranges centered at 315, 500, 800, 1250, 2000, and 3150 Hz with uniform distribution, and the noise bursts were filtered into the same 1/3 octave bands by 26th order Butterworth filters centered at the same frequencies. Empty 1/3 octave bands separated each of the bands to reduce spectral interactions between neighboring tones and noises.

Each cloud was 2 s in duration and was constructed independently so that the timing, tone frequencies, and tone levels were unique for every presentation. The number of tones and noises was set for the target and non-target bands in each condition and was equal in all non-target bands, resulting in a pre-defined number of tones and noises evenly distributed among the frequency bands. In conditions without noise bursts, there were an equivalent of thirty-six 100-ms tone units (including target constructions), equally distributed across the six frequency bands. In conditions with noise bursts, there were an equivalent of twenty 100-ms tone units and seven 100-ms noise units (including target constructions), distributed so that the five non-target bands were all equal, but the target band was slightly different because of the target constructions.

Each cloud was constructed by first randomly selecting the target frequency band and then randomly determining the different presentation frequencies for the tones in each band. After generating the tone and noise bursts for the cloud, a unique initial onset time was determined for the target band by randomly selecting a delay between 100 ms and 200 ms. Tones and noises in each band were uniquely ordered and separated by random lengths of silence. The longest possible length for within-band inter-burst silences was calculated as the total silence in the band divided by the number of events in that band, and the minimum length was one quarter of the maximum length. The target could occur at any time after the first 100 ms and before the last 100 ms of the interval. The levels of the distracting tones were set based on a Gaussian distribution with a mean of 45 dB SPL and standard deviation of 2 dB, and noise bursts were set at 75 dB SPL. Target tones were presented at 40 dB SPL.

The first condition (LT) required listeners to detect a long tone within a cloud of short tones. The second condition was the same, but the distracting clouds also included noise bursts (LTn). In the third condition (ILTn), listeners were asked to detect a long tone in mixtures of short tones and noise bursts, but in this case, no physical long tone was present; instead an illusory long tone was created by concatenating a short tone, a noise burst, and a second short tone at the same frequency to produce a tone-noise-tone sequence (onset and offset ramps overlapped at their half-amplitude points). The fourth condition (ST) required listeners to detect short tones in clouds of long tones, and the fifth condition (RTn) required the detection of a pair of repeated short tones (identical to the tone-noise-tone sequence, ILTn, but without the noise) in mixtures of short tones and noise.

Each experimental block consisted of 25 three-interval, three-alternative forced-choice (3I-3AFC) trials, and the task was always to identify which of the three 2-second intervals contained the designated target. The order of blocks was randomized for each listener. Each listener completed four blocks per condition, so that the performance of each listener was determined based on 100 trials per condition.

Twenty-eight young adult listeners (16 female; mean age of 23 years) with audiotically normal hearing were recruited from the University of Minnesota student population and provided written, informed consent. All experiments were approved by the University of Minnesota Institutional Review Board.

Results

Figure 2 shows the percentage of correctly identified intervals for each of 5 conditions, averaged across the 28 listeners; error bars indicate the standard error of the mean. A one-way repeated-measures ANOVA with the proportion of correct responses as the dependent variable found a significant main effect of condition, $F(4,108) = 86.3$, $p < 0.001$; $\eta^2 = 0.76$. The clearest measure of whether the illusion was successful is the difference in detection performance between the condition with the illusory long (ILTn) and the condition with the two repeated short tones (RTn): the only difference between these two conditions is the presence of the noise burst between the two short tones. A benefit of the continuity illusion was observed, with performance in the ILTn condition being significantly better than that in the RTn condition, according to a planned comparison paired t-test, $t(27) = 5.86$, $p < 0.0001$, $d = 1.27$.

A similar comparison of the LT and LTn conditions showed that replacing some masking tones with noise bursts slightly improved listeners' ability to detect the physical long tones embedded in the mixtures, $t(27) = -2.25$, $p = 0.033$; $d = -0.43$, perhaps because there were fewer masking tones, creating less scope for "confusion" between the noise bursts and the target tone than between the masking tones and target tone (Kidd, Mason, & Arbogast, 2002). A comparison of the LTn and ILTn conditions shows that the illusory long target tones (ILTn) led to a significantly poorer mean performance than the physical long tones (LTn), $t(27) = 7.0$, $p < 0.0001$; $d = 1.33$. Strong linear correlations between performance in the ILTn condition and performance in the LT and LTn conditions (LT and ILTn: Pearson's $r = 0.78$, $p < 0.0001$; LTn and ILTn: Pearson's $r = 0.73$, $p < 0.0001$) suggest that the ability

to detect real long tones within a mixture is a strong predictor of performance with the illusory long tones.

The results also confirm the perceptual asymmetry between long and short tones (Cusack & Carlyon, 2003) by showing better detection of both real and illusory long tones in a cloud of short tones than detection of a short tone in a cloud of long tones, LTn vs. ST: $t(27) = 12.77$, $p < 0.0001$, $d = 2.54$; ILTn vs. ST: $t(27) = 5.47$, $p < 0.0001$, $d = 1.12$.

Discussion

Our results indicate that illusory long tones formed by the continuity illusion are detectable even within complex masking conditions that impose a relatively high perceptual and attentional load. This outcome suggests that aspects of the illusion are formed early enough to permit illusory features to be mapped and used to extract objects from within complex auditory scenes. Nevertheless, despite the perceptual robustness of the illusion within the mixtures, a significant difference was observed between detection of real long tones and illusory long tones. The source of this difference is investigated further in Experiment 2.

Experiment 2

Rationale

In Experiment 1, illusory long tones (ILTn) were detected less reliably than real long tones (LT and LTn), which suggests either that the continuity illusion was not complete or that target detection was hampered in some other way in the ILTn condition. Experiment 2 was designed to test three hypotheses to elucidate the reasons underlying the difference in performance produced by the real and illusory tones.

The first hypothesis is that the continuity illusion was not complete, and that the two short tones, interspersed with noise, did not elicit as salient a percept of a long tone as the actual long tone. To test this possibility, direct comparisons were made between the real and illusory long tones, as well as between the illusory long tone and the repeated short tones separated by silence.

The second hypothesis is derived from the fact that the interrupting noise bursts in the ILTn condition were considerably higher in level than the target tones (75 dB vs. 40 dB SPL). A high level was used to ensure that an interruption between the two tones was not perceived, but the level difference may have also been sufficient for the noise to produce some forward or backward masking of the target tone (e.g., Oxenham & Moore, 1994). To test this possibility in Experiment 2, a new condition (ILTn-low) was designed as a repeat of ILTn, but with a lower noise level (55 dB SPL), which should still have been sufficient to induce the continuity illusion (e.g., Houtgast, 1974).

The third hypothesis to explain the lower average detection rate of the illusory long tones is that the strength of the continuity illusion may be variable among the listeners. It is possible that an individual who perceives the illusion weakly in isolation will also perceive it weakly within a complex masking background. Thus, the difference in performance observed in Experiment 1 would not be due to the complex background, but rather due to individual

differences in the strength of the illusion itself. To test this possibility, we measured the threshold noise level for each listener corresponding to the transition from hearing repeated tones with interrupting noise to hearing a continuous noise with a superimposed noise burst. This level has historically been referred to as the pulsation threshold (Duifhuis, 1980; Houtgast, 1974; Plack & Oxenham, 2000).

Methods

To test the first hypothesis, fourteen of the listeners from Experiment 1 completed a set of direct comparisons designed to determine the relative strength of the pop-out effect for the real long tones, the illusory long tones, and the repeated tones. The construction of tone clouds was identical to that in Experiment 1, but the clouds were presented in a 2-interval, 2-alternative forced-choice (2I-2AFC) paradigm. In all comparisons, the listener was required to identify which cloud interval contained the “best” or most obvious long tone. In the first condition, one cloud interval contained two short tones separated by a noise to induce the continuity illusion (ILTn), and the other contained just the two repeated tones (RTn). This condition was tested to confirm the conclusion from Experiment 1 that the continuity illusion results in a long tone percept that is more salient than a simple repeated tone in a background of random short tones. The second condition compared a cloud interval containing an illusory long tone (ILTn) with one containing a real long tone (LTn).

To test the second hypothesis, thirteen of the original normal-hearing listeners completed blocks consisting of presentations of clouds constructed as in Experiment 1, in a 3I-3AFC paradigm, but in a new condition, ILTn-low, where the level of the noise bursts within the tone clouds was set at 55 dB SPL instead of 75 dB SPL. The level of the target tones remained at 40 dB SPL.

To test the third hypothesis, the same thirteen listeners also completed an adaptive measurement of the noise level required for each listener to perceive an illusory long tone, given a physical tone-noise-tone sequence (pulsation threshold), with the tones presented at the same level as in Experiment 1 (40 dB SPL). At low noise levels, it was expected that listeners would not report hearing a long tone, as the gap between the two tones would be clearly audible; at high noise levels, the continuity illusion was expected to produce the percept of a long tone. The adaptive procedure was designed to determine the noise level corresponding to the transition between the two percepts. The same tone-noise-tone sequence of 100-ms components used throughout Experiments 1 and 2 was used but was presented in isolation, rather than in the clouds of tone and noise bursts used in the previous experiments. Thresholds were measured for tone frequencies of 315, 1000, and 3150 Hz, frequencies that span the range over which the target tone was presented in Experiment 1. Thresholds were measured using an interleaved, dual adaptive tracking procedure. In track 1, a 1-up, 2-down method was used, yielding the threshold for 70.1% long tone perception; in track 2, a 2-up, 1-down method was used, yielding the level where an individual perceived the long tone 29.9% of the time. The average of these two tracks is the noise level where the listener would be expected to perceive an illusory long tone 50% of the time (Jesteadt, 1980).

The starting noise levels in tracks 1 and 2 were 60 and 35 dB SPL, respectively. The step size of the tracking procedure in each track was initially 5 dB, and was reduced to 2 dB after two reversals. Thresholds were computed from the final 4 reversals of 8 reversals recorded in each track. The interleaved tracks were not balanced; each terminated independently after 8 reversals. Each listener completed three interleaved track pairs at each test frequency. Upper and lower tracks were averaged, and pulsation threshold measurements below 10 dB SPL (far below levels assumed necessary for formation of the illusion) were excluded. One listener was dropped entirely (all three tracks at 315 Hz and two tracks at 3150 Hz resulted in unrealistic threshold measurements), and one listener had one of three measurements dropped from the 315 Hz and 1000 Hz conditions. That listener was retained, but with data based on 2 measurements at those frequencies instead of 3. Differences were calculated between paired upper and lower tracks only for measurements that met the minimum pulsation threshold discussed above.

Results

In the direct comparison of tone-noise-tone intervals and repeated tone intervals, listeners selected the tone-noise-tone sequence over the repeated tone sequence about 80% of the time (see Figure 3). The distinct preference for the illusory long tone construction provides direct support for the indirect conclusion from Experiment 1 that an interrupting noise between two short tones produces the illusion of a long tone, even in the presence of a complex acoustic background, and that the noise-interrupted tones were perceived as more continuous than the repeated tones alone.

The direct comparisons of real long tone intervals with illusory long tone intervals resulted in listeners selecting the real long tone in an average of only about 50% of trials, which could be interpreted to mean that listeners were not able to discriminate reliably between the real and illusory long tones. Scores in the ILTn condition of Experiment 1 were correlated with percentage of illusory long tones selected in direct comparison intervals, Pearson's $r = 0.54$, $p = 0.046$, but the difference between LTn and ILTn performances in Experiment 1 was not significantly correlated with preference for illusory long tones in direct comparisons, Pearson's $r = -0.37$, $p = 0.20$.

Our second hypothesis for why the results of Experiment 1 indicate better detection of real long tones (LTn) than illusory long tones (ILTn) was that the high level of the noises in the mixture may have partially masked the tone components of the illusion. In Experiment 2 we included a condition with lower-level noise bursts (ILTn-low). A repeated-measures ANOVA comparing performance in the original condition with performance using the lower noise level (ILTn-low) revealed no significant effect of noise level, $F(1,12) = 3.1$, $p = 0.1$; $\eta^2 = 0.21$ (see Figure 4), suggesting that changing the noise burst level from 75 to 55 dB SPL did not significantly alter performance. This result suggests that lower performance in the ILTn condition compared to the LTn condition was not necessarily due to partial masking of the tones by the high-level noise. The possibility remains that a reduction of energetic masking may have been counterbalanced by a reduction of illusion strength. We believe this to be unlikely, as noises 15 dB above the level of the tones they interrupt can still be expected to produce a robust illusion.

The third hypothesis for the different detection rates found for real and illusory long tones in Experiment 1 was that perception of the long tone illusion may vary from listener to listener, which, in turn, may result in inter-individual differences in detection of the illusion within the clouds. Pulsation thresholds, corresponding to the noise level at which the illusion was reported 50% of time, were measured for signal frequencies of 315, 1000, and 3150 Hz. As mentioned above, one listener was not able to complete the task reliably, and those data were removed from our analyses. Mean pulsation thresholds across the remaining 12 listeners did not differ significantly across the three target frequencies, $F(2,22) = 1.89$; $p = 0.18$; $\eta^2 = 0.15$, and no (Pearson product moment) correlations between the pulsation thresholds and the rate of illusion detection in the ILTn condition of Experiment 1 were significant; 350 Hz: $r = 0.38$, $p = 0.23$; 1000 Hz: $r = 0.57$, $p = 0.054$; 3150 Hz: $r = 0.038$, $p = 0.23$. Similarly, the average pulsation thresholds (combined across frequencies) were not correlated with illusory long tone detection rate in the ILTn condition, $r = 0.56$, $p = 0.058$. Possibly more notable is that no significant Pearson's correlation was found between mean pulsation thresholds and the *difference* in target identification rate between the LTn and ILTn conditions in Experiment 1, $r = -0.35$, $p = 0.27$. Thus, the pulsation-threshold data do not support our third hypothesis, which was that lower susceptibility to the continuity illusion in isolation may result in poorer detection of the illusion in a complex mixture (condition ILTn) or more reduction in target identification between the LTn and ILTn conditions.

Some of the pulsation thresholds measured in this experiment were lower than would be expected in order to produce a strong continuous percept with 40 dB SPL tones. For instance, some of the measured thresholds were as low as 35 dB SPL, whereas many previous studies have found that the noise level should normally exceed that of the tone by at least a few dB. This outcome suggests that some listeners may have had unusually low criteria for labeling a presentation as continuous. As suggested in Vinnik et al. (2011), variations in subjective understanding of the task and what constitutes a “repeated” or “continuous” tone may impact how threshold measurements relate to illusory target perception in the mixture. We quantified the degree of uncertainty in threshold measurements by calculating the difference between the thresholds obtained from the upper and lower tracks in the interleaved adaptive tracking procedure. This difference has been proposed as a measure of sensitivity, or just-noticeable difference (Jesteadt, 1980). More reliable judgments should result in smaller differences between tracks, whereas more variable judgments should typically result in larger differences.

The difference measure was found to correlate significantly and negatively with percent correct in the detection of illusory long tones in the ILTn condition of Experiment 1, Pearson's $r = -0.67$, $p = 0.016$. In other words, subjects who were able to more reliably judge perceptual continuity in isolation were more likely to detect the illusory tone in a complex mixture. In contrast, however, there was no significant correlation between the threshold difference measure and the *difference* in target identification between the LTn and ILTn conditions, Pearson's $r = -0.17$, $p = 0.59$. Thus, contrary to the hypothesis, a more reliable illusory percept did not predict more similar performance between real and illusory tones in a complex background. This pattern of results may be explained by the additional

finding that there was also a significant correlation between the threshold difference measure and performance in the real long tone condition (LTn) of Experiment 1, Pearson's $r = -0.75$, $p = 0.004$. In other words, the subjects who were more reliable in reporting the threshold for the continuity illusion were also better performers in detecting both the real and the illusory long tones in the complex background.

Discussion

Experiment 2 was designed to test three hypotheses to explain the result from Experiment 1 that the illusory long tone was detected less readily than the real long tone within a cloud of short-tone distractors.

The first hypothesis, that the illusory tone resulted in a less salient sensation than the real long tone, was not supported: listeners consistently selected the illusory long tone over pairs of repeated short tones but, on average, did not consistently select the real long tone over the illusory long tone when each was presented within the complex acoustic background. This outcome is consistent with the idea that the continuity illusion was successfully generated, but it seems inconsistent with results from Experiment 1, which showed significantly poorer performance for illusory long tones (ILTn) than for real long tones (LTn). This apparent discrepancy may be due to differences in the designs of the two tests. In the current experiment, one real and one illusory long tone were presented in each trial. Given the results from Experiment 1, it is likely that in some trials at least one of the two long tones was not detected. One can assume that in cases where only one tone was detected, listeners would select the interval containing the detected tone, whereas in cases where neither tone was detected, responses should be distributed roughly evenly (50%) between the two intervals. The final possibility involves trials where both tones were detected. Responses there may have been ambiguous, and subject-dependent, with some subjects preferentially selecting the real long tone and other selecting the illusory tone. Simple simulations of detection probabilities, based on performance in Experiment 1, show that performance in selecting the real long tone over the illusory long tone can be predicted to lie anywhere from about 30% to 90%, depending on the assumed preference for real over illusory tones. Note that this intuitive account can predict 50% performance without assuming that the real and illusory tones are indistinguishable.¹ Thus, the results of Experiment 2, that listeners consistently selected the illusory long tone over pairs of repeated short tones, provide further evidence that the continuity illusion occurs even in complex acoustic scenes and that it can be used to induce a perceptual “pop-out” of the target tone. These data did not, however, clarify whether or under what conditions real long tones pop out of complex mixtures more readily than do illusory long tones.

The second hypothesis, that the noise level used in Experiment 1 may have resulted in partial forward or backward masking of the target tone, was also not supported: results using a lower-level masker produced very similar results, suggesting that partial masking by the noise did not limit performance.

¹Although the intuitive explanation provided here for performance in Experiment 2 relies on a “high-threshold” account of signal detection, a similar outcome can be derived using the more usual continuous distributions assumed under Signal Detection Theory (Green & Swets, 1966).

The third hypothesis, that detection is affected by individual differences in the salience of the continuity illusion, was also not supported. No relationship was found between the pulsation threshold (the noise level above which the tone is generally perceived as continuous) and either performance in the ILTn condition from Experiment 1 or the real long tone advantage (difference between LTn and ILTn). Listeners' real long tone advantage was similarly not correlated with a measure of the reliability of the illusory percept (quantified by the difference between two adaptive tracks measuring the 29% and 71% points on the psychometric pulsation threshold function). However, the ability to detect both the illusory and real tone was significantly correlated with pulsation threshold reliability, suggesting that better performance in Experiment 1 correlated with more reliable performance in Experiment 2, but not suggesting any link between the strength of the illusion with and without the presence of background sounds.

In summary, none of the three hypotheses tested in Experiment 2 were supported by the data. This leaves the source of the difference in detection between LTn and ILTn conditions somewhat unclear. It may be simply that the two tones interspersed by noise were sufficient to induce the illusion of continuity, but did not produce quite as salient a long tone as the actual long tone, perhaps in part because the physical duration of the tonal part of the illusory tone was only 2/3 as long as the physical duration of the actual long tone.

General Discussion

This study provides behavioral evidence that the continuity illusion persists even in complex and unpredictable acoustic environments, and that it can be exploited to form perceptual objects and improve detection performance, even without a priori knowledge of the target's spectrotemporal location.

Our task was derived from investigations carried out by Cusack and Carlyon (2003), who reported an asymmetry between the detection of long tones in a background of short tones and the detection of short tones in a background of long tones. They suggested that this perceptual asymmetry may be derived from the relatively greater population of duration-sensitive neurons tuned to long versus short tones. Carlyon et al. (2009) went on to study the duration illusion and its parallels to the continuity illusion. They hypothesized that the continuity illusion, like the duration illusion, is based on ambiguity of onsets and offsets, ambiguity that may be enhanced by adding competing tones and noises to the presentation to draw perceptual resources away from the onsets and offsets of the target construction (Kobayashi & Kashino, 2012). According to this line of reasoning, the presence of competing tones and noises in our paradigm may not impair the formation of the illusory long tone percept, and may even support it by drawing perceptual resources away from the detection of onsets and offsets in the tone-noise-tone construction.

We found that the perceptual asymmetry reported by Cusack and Carlyon (2003) remains when long tones are induced via illusion, albeit to a lesser extent than when real long tones are used. This suggests that the continuity illusion is established prior to feature mapping for scene analysis and detection of target objects, although the formation of the illusion may not be complete.

Our experimental paradigm attempted to minimize peripheral auditory factors by ensuring that there was always at least a 1/3-octave separation between neighboring spectral components within the acoustic stimulus. Thus, it appears likely that auditory scene analysis in these tasks is limited at relatively higher levels of the auditory system. Neuroimaging results, in combination with behavior, have suggested that the neural correlates of detecting a sequence of target tones in a cloud of equal-duration masking tones emerges in auditory cortex (Wiegand & Gutschalk, 2012). Ongoing combined neuroimaging and behavioral techniques may help to establish the neural correlates and potential locus of both duration-dependent target detection and the continuity illusion in complex acoustic backgrounds.

Acknowledgments

Supported by NIH grant R01 DC007657.

References

- Bashford JA, Riener KR, Warren RM. Increasing the intelligibility of speech through multiple phonemic restorations. *Percept Psychophys*. 1992; 51(3):211–217. [PubMed: 1561046]
- Bregman, A. Auditory Scene Analysis: The Perceptual Organization of Sound. Cambridge, MA: MIT Press; 1990.
- Carlyon RP, Deeks JM, Shtyrov Y, Grahn J, Gockel HE, Hauk O, Pulvermüller F. Changes in the perceived duration of a narrowband sound induced by a preceding stimulus. *J Exper Psychol Hum Percept Perform*. 2009; 35(6):1898–1912. [PubMed: 19968442]
- Cusack R, Carlyon RP. Perceptual asymmetries in audition. *J Exper Psychol Hum Percept Perform*. 2003; 29(3):713–725. [PubMed: 12848335]
- Duifhuis H. Level effects in psychophysical two-tone suppression. *J Acoust Soc Am*. 1980; 67(3):914–927. [PubMed: 7358916]
- Green, DM.; Swets, JA. Signal Detection Theory and Psychophysics. New York: Wiley; 1966.
- Gutschalk A, Micheyl C, Oxenham AJ. Neural correlates of auditory perceptual awareness under informational masking. *PLoS Biol*. 2008; 6(6):e138. [PubMed: 18547141]
- Haywood NR, Chang ICJ, Ciocca V. Perceived tonal continuity through two noise bursts separated by silence. *J Acoust Soc Am*. 2011; 130(3):1503–1514. [PubMed: 21895090]
- Heinrich A, Carlyon RP, Davis MH, Johnsrude IS. The continuity illusion does not depend on attentional state: fMRI evidence from illusory vowels. *J Cog Neurosci*. 2011; 23(10):2675–2689.
- Houtgast T. Psychophysical evidence for lateral inhibition in hearing. *J Acoust Soc Am*. 1972; 51(6):1885–1894. [PubMed: 4339849]
- Houtgast T. Lateral suppression and loudness reduction of a tone in noise. *Acustica*. 1974; 30(4):214–221.
- Jesteadt W. An adaptive procedure for subjective judgments. *Percept Psychophys*. 1980; 28(1):85–88. [PubMed: 7413416]
- Jones DM, Macken WJ, Murray aC. Disruption of visual short-term memory by changing-state auditory stimuli: the role of segmentation. *Mem Cog*. 1993; 21(3):318–328.
- Kidd G, Mason CR, Arbogast TL. Similarity, uncertainty, and masking in the identification of nonspeech auditory patterns. *J Acoust Soc Am*. 2002; 111(3):1367–1376. [PubMed: 11931314]
- King AJ. Auditory Neuroscience: Filling in the Gaps. *Curr Biol*. 2007; 17(18):R799–801. [PubMed: 17878047]
- Kitagawa N, Igarashi Y, Kashino M. The tactile continuity illusion. *J Exper Psychol*. 2009; 35(6):1784–1790.
- Kobayashi M, Kashino M. Effect of flanking sounds on the auditory continuity illusion. *PloS One*. 2012; 7(12):e51969. [PubMed: 23251666]

- Komatsu H. The neural mechanisms of perceptual filling-in. *Nat Rev Neurosci.* 2006; 7(3):220–231. [PubMed: 16495943]
- Micheyl C, Carlyon RP, Shtyrov Y, Hauk O, Dodson T, Pullvermüller F. The neurophysiological basis of the auditory continuity illusion: A mismatch negativity study. *J Cog Neurosci.* 2003; 15(5): 747–758.
- Miller G, Licklider J. The intelligibility of interrupted speech. *J Acoust Soc Am.* 1950; 22:167–173.
- Oxenham AJ, Moore BCJ. Modeling the additivity of nonsimultaneous masking. *Hear Res.* 1994; 80(1):105–118. [PubMed: 7852196]
- Petkov C, Sutter M. Evolutionary conservation and neuronal mechanisms of auditory perceptual restoration. *Hear Res.* 2011; 271:54–65. [PubMed: 20541597]
- Plack CJ, Oxenham AJ. Basilar-membrane nonlinearity estimated by pulsation threshold. *J Acoust Soc Am.* 2000; 107(1):501–507. [PubMed: 10641658]
- Riecke L, Micheyl C, Oxenham AJ. Global not local masker features govern the auditory continuity illusion. *J Neuroscience.* 2012; 32(13):4660–4664.
- Riecke L, van Opstal AJ, Goebel R, Formisano E. Hearing illusory sounds in noise: sensory-perceptual transformations in primary auditory cortex. *J Neuroscience.* 2007; 27(46):12684–12689.
- Riecke L, Vanbussel M, Hausfeld L, Ba Kent D, Formisano E, Esposito F. Hearing an illusory vowel in noise: suppression of auditory cortical activity. *J Neuroscience.* 2012; 32(23):8024–8034.
- Shahin AJ, Bishop CW, Miller LM. Neural mechanisms for illusory filling-in of degraded speech. *NeuroImage.* 2009; 44(3):1133–1143. [PubMed: 18977448]
- Shinn-Cunningham BG, Wang D. Influences of auditory object formation on phonemic restoration. *J Acoust Soc Am.* 2008; 123(1):295–301. [PubMed: 18177159]
- Treisman AM, Gelade G. A Feature-Integration Theory of Attention. *Cog Psychol.* 1980; 12:97–136.
- Vinnik E, Itskov PM, Balaban E. Individual differences in sound-in-noise perception are related to the strength of short-latency neural responses to noise. *PLoS One.* 2011; 6(2):e17266. [PubMed: 21387016]
- Warren, R. *Auditory Perception: A New Analysis and Synthesis.* Cambridge, UK: Cambridge University Press; 1999.
- Warren R, Obusek C, Ackroff J. Auditory induction: perceptual synthesis of absent sounds. *Science.* 1972; 176:1149–1151. [PubMed: 5035477]
- Wiegand K, Gutschalk A. Correlates of perceptual awareness in human primary auditory cortex revealed by an informational masking experiment. *NeuroImage.* 2012; 61(1):62–69. [PubMed: 22406354]

Figure 1.

Schematic diagram of the five conditions tested. Random-frequency distractors, or “clouds,” in Conditions 1 (LT) and 2 (LTn) consisted of short tones and, in the case of Condition 2, short noise bursts; the target tone in these clouds was a physical long tone. Clouds in Condition 3 (ILTn) were made of short tones and noise, and targets were illusory long tones elicited by inserting a noise between two same-frequency short tones. In Condition 4 (ST), a short target tone was imbedded in a cloud of long masking tones. The target in Condition 5 (RTn) was a tone-noise-tone sequence with the noise removed, resulting in a repeated short tone; the stimulus parameters otherwise matched Conditions 2 and 3.

Figure 2.

Each bar depicts the mean percentage and standard error of target intervals correctly identified by 28 listeners. The LT and LTn conditions produced a relatively high rate of real long tone detection, compared to the significantly lower rate of illusory long tone detection in the ILTn condition. Detection of short (ST) and repeated tones (RTn) was significantly poorer than detection of real or illusory long tones.

Figure 3.

In direct comparisons, the illusory tone was clearly preferred over the repeated tones (left bar). There was, however, no clear preference for real over illusory long tones (right bar). Standard errors are indicated.

Figure 4.

Results from the 13 listeners who completed both the ILTn and ILTn-low conditions. In the ILTn-low condition the level of the noise in the masking cloud and in the tone-noise-tone sequence was lowered to 55 dB instead of 75 dB SPL (the noise level in all other conditions).