BINAURAL FACTORS IN AUDITORY CONTINUITY

C.J.Darwin, M.A.Akeroyd & R.W. Hukin

Experimental Psychology, University of Sussex
Brighton BN1 9QG, UK
cjd@biols.susx.ac.uk

ABSTRACT

Auditory continuity is a powerful illusion which has implications for the efficient coding of sound. Although auditory continuity has been studied extensively as a monaural phenomenon, there have been few reports of the influence of binaural factors. Hartmann [1] and Kashino & Warren [2] report that when a tone alternates with a noise, the continuity threshold of the tone is lower when the tone and noise differ in their interaural timing. Here we aconfirm the binaural contribution to continuity continuity using Huggins pitch [3]. We alternate a Huggins-pitch noise (target) with a negative-phase Schroeder complex (inducer) and find that pulsation thresholds are lower when listeners judge continuity of the pitch percept within the Huggins noise, than when they judge continuity of the noise itself. We also demonstrate that binaural continuity is determined by the relative ITDs of the targe and inducer rather than by their perceived spatial positions and introduce a simple model to account for our findings, which assumes that the tone is heard as pulsating if there is a detectable dip in the output of either of three channels: two monaural [4] and one binaural channel.

1. INTRODUCTION

Most sounds that we hear are partially masked in frequency and time by other sounds. Yet neither our perceptual experience, nor our ability to recognise sounds, is unduly disturbed by such masking. In the laboratory, this can be demonstrated by two related illusions. First, we experience as continuous a sound that is briefly replaced by an appropriately more-intense sound [5-7]. For example, noise bursts alternating with brief periods of silence are heard as pulsing, but turn into a continuous noise when the silences between them are replaced by a more intense noise of similar spectrum. The louder noise induces continuity of the quieter noise. A similar illusion occurs when gaps between tone pulses are filled with a loud enough noise. The two effects have been termed respectively “homophonic induction” and “heterophonic induction" by Warren [8], since in the former case the spectral composition of the two sounds are similar, but in the latter they are radically different. Second, we perceptually restore predictable speech segments that are replaced by an appropriate noise [9, 10]. Two related effects also occur for speech stimuli with parts repeatedly removed and masking noise inserted: there is an impression of continuity for the speech [5] and, for redundant speech, there is also an improvement in intelligibility [11, 12]. The general explanatory principle behind these “auditory continuity” and perceptual restoration phenomena is that they occur *when the interrupting noise would have masked the deleted sound had it actually been present.* In other words they occur when the listener lacks information that the deleted sound is actually absent.

A considerable amount of research has investigated both of these phenomena monaurally [for recent examples see 13, 14]. Little attention, however, has been paid to their interaction with binaural processes concerned with determining the spatial locations of sounds or detecting sounds in adverse environments.

Only three papers have looked at auditory continuity with binaural stimulation. Hartmann [1] and Kashino & Warren [2] measured continuity thresholds (or equivalently pulsation thresholds) for tones alternating with diotic noise (heterophonic induction). Both papers report that when low-frequency tones were presented with reversed polarity in one ear (NoSπ) continuity thresholds were lower by about 6 dB than when the polarity was not reversed (NoSo). This result indicates that the masking noise was less effective in the NoSπ than in the NoSo. Both authors conclude that effects similar to (though perhaps smaller than) the binaural masking-level difference can be obtained with continuity thresholds. No mechanism has been proposed for this effect. In particular it is not clear whether the effect is due to the different phase relations of the stimuli per se or to the sounds' different locations. One’s subjective impression of NoSπ is of the diotic noise in the middle of the head with the reversed polarity continuous tone coming from one side. Kashino & Warren also reversed the polarity of the noise, rather than the tone, (NπSo) and found a similar change in continuity threshold. Here the location of the noise is diffuse about the midline while the the tone is compactly localised on the mid-line.

Thurlow [4], by contrast, manipulated interaural intensity relations between alternating pairs of diotic white noise bursts (homophonic induction). On the basis of his results, he argued that, a binaurally-presented sound will only be heard as *dis*continuous if it can be heard as *dis*continuous in either ear alone. Thurlow thus proposes a monaural condition for binaural continuity. However, the results of the other two papers suggests that Thurlow's monaural condition for continuity (if it generalises from homophonic to heterophonic induction) may be necessary, but not sufficient: a binaurally reversed polarity tone in diotic noise (NoSπ) which is heard as continuous in each ear separately, may be heard as discontinuous binaurally.

The first experiment that we report here confirms the binaural contribution to auditory continuity using Huggins pitch [3], a binaural white noise that generates a weak pitch sensation when presented binaurally, but not monaurally. We alternate a Huggins-pitch noise with a negative-Schroeder-phase complex [15] - a sound with a very different timbre - and ask listeners to judge the continuity either of the Huggins noise itself, or of the pitch generated by the Huggins noise. In keeping with the conclusions of Kashino and Warren, we demonstrate that the Huggins stimulus needs to be quieter to give continuity of its (lateralised) pitch percept than of the (centered) percept of the noise itself.

The second experiment uses sounds similar to those used by Kashino & Warren, but additionally manipulates interaural instensity differences to ask whether the decreased continuity threshold produced by an interaural time difference is due to the ITD per se, or to the different perceived spatial positions produced by the ITDs.

Finally we introduce a schematic binaural model that can account qualitatively for the results.

2. EXPERIMENT 1

In this experiment we confirm that there is a binaural contribution to auditory continuity by asking listeners to judge the perceived continuity of a pitch that can only be heard as a result of binaural interaction. We ask whether the threshold for continuity of this pitch is lower than that for continuity of the noise that generates the pitch.

2.1. Stimuli & Procedure

On each trial listeners heard eight alternations of two 120-ms sounds, the target and the inducer, with overlapping 10-ms rise-fall times. The target was a Huggins-pitch stimulus, the inducer was a negative-Schroeder-phase complex, which acted as the inducer of continuity for the other sound.

For the Huggins-pitch stimulus one ear of the listener received a 300-Hz to 1-kHz band-pass filtered white noise; the other received simultaneously the same noise, but with a one ERB wide (77 Hz) linear phase change from 0 to 2π around 500 Hz. When the signal to either ear is heard alone the percept is simply of a band-pass noise; but both ears together gives the percept of a weak pitch at 500 Hz to one side, with a band-pass noise in the center of the head.

The inducing sound was a negative-Schroeder-phase complex. Such sounds are designed to have a maximally flat waveform envelope and were used to provide good induction of auditory continuity for the Huggins stimulus with a very different timbre, without the risk of clipping of extreme noise samples. The complex consisted of an equal-amplitude harmonic series with a fundamental of 60 Hz. The phase of the *ith* harmonic was $-\pi i(i-1)/n$, and the resulting complex was band-pass filtered (Hann-filter with 100-Hz skirts) between 300 and 1400 Hz. The resultant sound was played diotically and gave the percept of a low-frequency buzz in the center of the head.

To calibrate the relative levels of the two sounds, they were first filtered in a way that modelled the cochlear response at 500 Hz (4^{th}-order gammatone filter with center frequency 500 Hz and bandwidth 77 Hz). The relative levels of the two sounds were then adjusted to give equal energy through the filter in the 0-dB condition and the Huggins stimulus then attenuated in 1 dB steps to –12 dB to give a total of 13 stimuli. Seven audiometrically-normal listeners, experienced in psychoacoustic tests, were tested twice on a random sequence of 10 examples of the sounds. One test used attenuations 0 through -8 dB and they were asked to judge on each trial whether the noise (as distinct from the buzz) sounded continuous or pulsing. The other test used all the stimuli and listeners were asked to judge whether the weak pitch in the noise sounded continuous or pulsing.

2.2. Results

Figure 1 shows the average percent of trials on which listeners heard the Huggins noise (squares) or the Huggins pitch (triangles) as pulsing rather than continuous as a function of the level of the noise relative to the inducing complex. The data from individual listeners were fitted with an inverse tanh function to find the attenuation level on which they perceived continuity on 50% of the trials in each test (the pulsation threshold). The means of these levels were -4.5 dB and -7.0 dB for the noise judgments and for the speech judgments respectively giving a mean difference of 2.4 (± 0.58 s.e.m.) dB; a difference which gives $t_6 = 4.2$, $p < 0.005$ on a 1-tailed test.

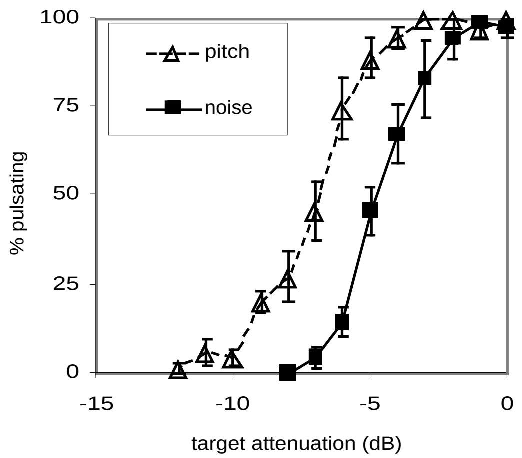

Figure 1. *Percent of trials on which listeners heard the Huggins noise or the Huggins pitch as pulsing rather than continuous as a function of the level of the noise relative to the inducing complex.*

2.3. Discussion

This experiment has demonstrated a binaural contribution to auditory continuity using the Huggins pitch - a percept which only arises binaurally. The pitch sensation is inaudible at either ear alone, and only arises through binaural interaction. The binaural contribution to continuity is directly implicated by the fact that the Huggins stimulus has to be 2.4 dB quieter (relative to the inducing complex sound) for listeners to hear the Huggins pitch as continuous than for them to hear the Huggins noise itself as continuous. This observation complements those made by Kashino and Warren and by Hartmann who varied the interaural phase relations of tones in noise in establishing a binaural contribution to auditory continuity.

None of these experiments however determine whether nature of the binaural contribution. Specifically, it is important to know whether the contribution arises at a low-level where different interaural cues to spatial location are dealt with separately, or at a higher level where they are combined to give a subjective location to a sound.

3. EXPERIMENT 2

This second experiment separates a binaural contribution to auditory continuity that is due to subjective spatial location from one that is due simply to the relative interaural time-differences of the target and inducer sounds. We make this separation by using individually-determined interaural level differences to return to the mid-line a sound lateralised by ITD.

3.1. Method

We first determined for each of our listeners the interaural level difference needed to return to their subjective midline a 120-ms 500-Hz tone that had an ITD of ±300 µs. The average level needed was about 8 dB and so this level will be used to refer to conditions that had the ILD.

We then measured the pulsation threshold for the 500-Hz tone, alternating with the same binaural negative-phase Schroeder complex as used in Experiment 1, for a variety of ITD/ILD combinations shown in the table:

ITD (µs)	ILD (dB)	Pulsation Threshold (dB)
300	0	-5.4
0	***0***	***-3.8***
-300	0	-5.2
300	***-8***	***-3.6***
0	-8	-2.0
-300	***8***	***-3.6***
0	8	-1.3

Entries in ***bold italic*** represent conditions where the tone was perceived (along with the continuity-inducing noise) in the midline. The design of the experiment allows us to compare the pulsation threshold for pairs of conditions with the same ILD, but which differ in ITD. If continuity is influenced by spatial location, then higher pulsation thresholds should be obtained for the ***bold italic*** conditions, irrespective of their ITD. However, if continuity is determined by ITD, the higher (less negative) continuity thresholds should be obtained in conditions that have 0 µs ITD rather than ±300 µs ITD.

3.2. Results

Illustrative identification functions for the upper and the lower pairs of conditions in the table are shown in Figure 2. In both the upper and the lower panels the lower pulsation threshold occurs for the condition where the tone has ±300 µs ITD, even though in the lower panel that condition has the tone heard (along with the inducing noise) in the midline. The pulsation thresholds for all the conditions are shown in the table. In all cases the pulsation threshold is lower for sounds that have ±300µs ITD rather than 0 µs ITD. Pulsation threshold thus varies with ITD, *not* with subjective location.

3.3. Discussion

Experiment 2 has shown that the binaural contribution to auditory continuity is determined by ITD rather than by subjective location. Models of auditory continuity might then be based on existing models of other phenomena related to the processing of ITD. We now present an outline of a theory of binaural continuity based loosely on Durlach's EC model of binaural processing [16].

Let us assume, following Durlach, that a central decision-making process has access to the output of auditory filters from each ear separately, and also to the output of a third binaural channel representing the cross-product of the signals from corresponding auditory filters at the two ears (after level equalization). Let us also assume that with monaural presentation listeners hear the target sound as continuous if there is no energy dip in the output of relevant auditory filters during the inducing sound. For binaural sounds let us also finally assume that listeners only hear the sound as continuous provided there is no energy dip in both of the monaural channels, and in the binaural channel. Processing the sounds that we have used through a binaural model incorporating these features [17] gives at least qualitative agreement with the results that we have obtained; specifically it predicts lower pulsation thresholds for the conditions that have ±300 µs ITDs than for those that have 0µs ITDs.

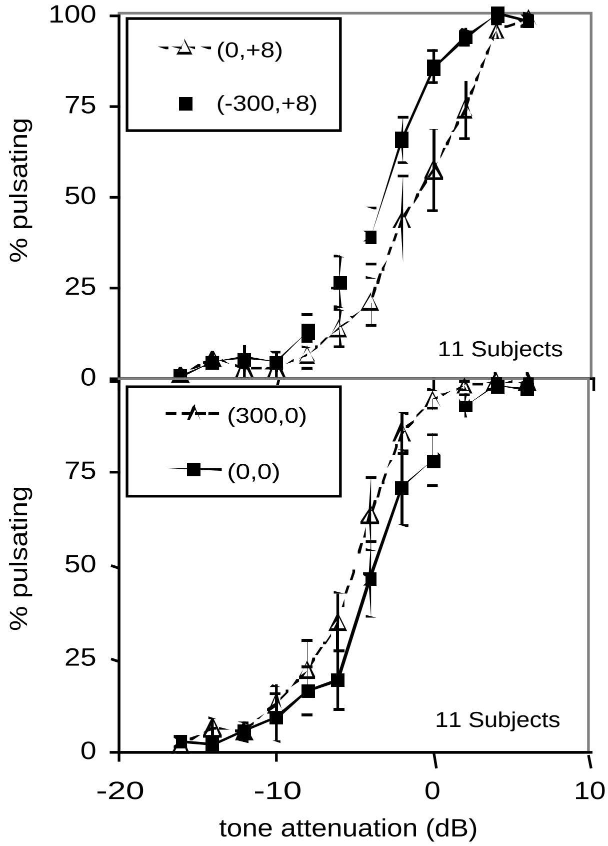

Figure 2. *Percent of trials on which listeners heard the tone as pulsating rather than continuous as a function of the level of the tone relative to the inducing noise*

4. CONCLUSIONS

The two experiments described here have confirmed a binaural contribution to auditory continuity using Huggins-pitch sounds, and have demonstrated that this binaural contribution arises at the level of processes that are sensitive to ITD rather than spatial position. We have sketched an explanatory model for the effects that we have found which will make a binaural sound continuous provided that there are no energy dips during the inducing sound in either of two monaural channels, or in a binaural channel responding to the cross-product of the two monaural channels.

5. ACKNOWLEDGEMENTS

Michael Akeroyd was supported by an MRC Career Development Award, and Rob Hukin by a grant from the UK EPSRC.

6. REFERENCES

[1] W. M. Hartmann, "A search for lateral inhibition," *Journal of the Acoustical Society of America*, vol. 75, pp. 528-535, 1984.
[2] M. Kashino and R. M. Warren, "Binaural release from temporal induction," *Perception & Psychophysics,* vol. 58, pp. 899-905, 1996.
[3] E. M. Cramer and W. H. Huggins, "Creation of pitch through binaural interaction," *Journal of the Acoustical Society of America*, vol. 30, pp. 413-417, 1958.
[4] W. R. Thurlow, "Auditory continuity effects with binaural stimuli," *Perception & Psychophysics*, vol. 42, pp. 173-9., 1987.
[5] G. A. Miller and J. C. R. Licklider, "The intelligibility of interrupted speech," *Journal of the Acoustical Society of America*, vol. 22, pp. 167-173, 1950.
[6] G. Vicario, "L'effetto tunnel acustico.," *Riv. di Psicol.*, vol. 54, pp. 41-52, 1960.
[7] W. R. Thurlow, "An auditory figure-ground effect," *American Journal of Psychology*, vol. 70, pp. 653-654, 1957.
[8] R. M. Warren, "Perceptual restoration of obliterated sounds," *Psychological Bulletin*, vol. 96, pp. 371-383, 1984.
[9] R. M. Warren, "Perceptual restoration of missing phonemes," *Science*, vol. 167, pp. 392-393, 1970.
[10] A. G. Samuel, "Phonemic restoration: insights from a new methodology," *J Exp Psychol Gen*, vol. 110, pp. 474-94., 1981.
[11] G. L. Powers and J. C. Wilcox, "Intelligibility of temporally-interrupted speech with and without intervening noise," *Journal of the Acoustical Society of America*, vol. 61, pp. 195-199, 1977.
[12] J. Verschuure and M. P. Brocaar, "Intelligibility of interrupted meaningful and nonsense speech with and without intervening noise," *Perception & Psychophysics*, vol. 33, pp. 232-240, 1983.
[13] R. P. Carlyon, J. Deeks, D. Norris, and S. Butterfield, "The Continuity Illusion and Vowel Identification," *Acustica*, in press.
[14] C. J. Plack and L. J. White, "Perceived continuity and pitch perception," *Journal of the Acoustical Society of America*, vol. 108, pp. 1162-9., 2000.
[15] M. R. Schroeder, "Synthesis of low peak-factor signals and binary sequences with low autocorrelation.," *IEEE Trans. Inf. Theory,*, vol. 16, pp. 85-89, 1970.
[16] N. I. Durlach, "Equalization and cancellation theory of binaural masking level differences," *Journal of the Acoustical Society of America*, vol. 35, pp. 1206-1208, 1963.
[17] M. A. Akeroyd, "A binaural cross-correlogram toolbox for MATLAB," 2001. www.biols.susx.ac.uk/home/Michael_Akeroyd/