COMPOSING WITH ABSENT SOUND

Christopher Haworth
Sonic Arts Research Centre
Queens University Belfast
BT7 1NN
chaworth01@qub.ac.uk

ABSTRACT
This paper describes a process of ‘composing with absent sound’ in the realisation of Correlation Number One, an 8-channel computer music piece that uses high frequency tones to generate distortion-product otoacoustic emissions (DPOAE) in the listener’s ears. Following a survey of other works created using the phenomena, the artistic motivations and technical implementation of the piece are described. Through a reflection on form arising as a product of limitation - in this case, that of the human auditory system - a notion of the ear as an active, productive sense organ emerges, in contrast to contemporary sound studies discourse that favours a depiction of the ear as a passive receiver.

1. INTRODUCTION
In the literature on sonic art and sound experience we often come across the notion of the ear as the submissive sense organ, ever open and vulnerable, its mechanical function (hearing) continuing indefinitely after its active, critical function (listening) has disengaged. ‘We have no ear lids’ [1]; ‘there is shut-eye, but no shut-ear’ [2]; ‘the ear never closes’ [3]. Implied here is the idea of the ear as a victim, defenceless against an onslaught of sound from all sides. No doubt when we consider extremes, such as the music of Merzbow or the diffuse, ever-present spread of muzak across public places, this latent potential of everyday sonic experience is brought to our attention: we become aware of our susceptibility to potential of everyday sonic experience is brought to our present.

But the ear is no mere passive receiver. It is a transducer, receiving acoustic energy and transforming it into electrical nerve impulses. Furthermore, as with other conversion mechanisms the relationship between the ‘input’ and the ‘output’ is not always linear, and it is from this interruption that the sciences of acoustics and psychoacoustics – respectively, the study of physical sound and the hearing mechanisms’ transduction of psychoacoustics – are derived.

This ‘gap’, between sound ‘out there’ and sound ‘in here’, can be seen as an obstacle to the composer, who wants his work to be ‘unbreakable’ from conception to reception. Common questions we encounter such as, ‘why, when I place a sound down there does it sound like it’s been placed up here?’ arise when the gap appears most wide - put simply, it happens when reality disappoints the ideality of the imagination. Perhaps moreso than in any other discipline, it is the artist who creates electronic music that must grapple with the science of perception in order to realise his work.

Computers offer such subtle control that the questions one finds oneself asking may be equal parts psychophysical and compositional. Yet the gap need not be viewed as an obstacle: looked at differently, the situation presents an opportunity. To this end, responses to the question: ‘what is the minimum amount a tone must differ in frequency such that it may be perceptible to listeners?’ may be found, not only within the many studies of frequency-difference limens by Brian Moore et al [4, 5, 6], but also within Alvin Lucier’s compositions In Memorium John Higgins (1982) and Crossings (1984), which are based around near-imperceptibly slow sinewaves sweeps. It is the same kind of questioning that led me to study perceptual idiosyncracies and psychoacoustic effects as potential compositional material.

2. PREVIOUS WORK
A previous work entitled The Law of the First Wavefront (2008) uses only the precedence effect to spatialize sound. There is no panning in the piece; the spatialization is produced via the manipulation of slight (< 2ms) time delays between identical sound grains presented in pairs of speakers. Since the auditory system suppresses late arriving reflections occurring up to 35ms after a direct sound is heard,1 this means that a) we hear a single ‘fused’ sound despite their being two sounds arriving at slightly different times, and b) the fused sound is perceived as coming from only the speaker projecting the earliest sound, despite sound being present in both. By gradually changing the rate at which one of the sounds is projected (either with a variable delay or with pitch shifting), as the two channels go further out of sync, we hear a single sound stream moving towards the speaker projecting the leading wavefront and away from the speaker projecting the lag.

This form of spatialization produces a different effect to other methods and is certainly worthy of further exploration. However, I found that the piece does not carry its primary concept - of time and space as interdependent, subjectively produced categories - as well as I had hoped, since only expert listeners seem to be able to tell they are hearing a ‘non-standard’ form of spatialization. It was this dissatisfaction that led me to investigate ‘distortion product oto-acoustic emissions’ (DPOAE), a much more commonly used phenomena in artistic practice.

---

1 The echo threshold is entirely dependant upon source characteristics.
3. BACKGROUND

DPOAEs are one example of auditory nonlinearity, a concept first suggested by Helmholtz [7]. They are audible ‘third tones’ generated in response to closely spaced pairs of pure tones with a ratio between 1.1 and 1.3, and which ‘activate the cochlea in the same region of the basilar membrane’ [8]. Of all psychoacoustic phenomena, DPOAEs have perhaps the most ‘musical’ history. Within musicology they bear the name ‘Tartini tones’ after the composer Giuseppe Tartini, who noticed the effect could be used to estimate the tuning and intonation of double-stops for the violin [9]. As such, a large number of composers and sonic artists have sought to use the phenomena in their work, notably Maryanne Amacher, Phill Niblock and most recently, Jacob Kierkegaard (as a recent ICMC paper surveyed [10]). In each of these cases the DPOAE is used differently. In Niblock’s work the effect is mostly subliminal - they appear as a product of the high intensity sound his music demands, but this means they can easily be mistaken for acoustic vibrations due to the density of the forms he creates. Amacher’s work utilises very high volumes too, but we are very much aware that we are being ‘played’, so to speak: that our ears are both receiving and producing the architecture of the piece. In such pieces as ‘Head Rhythm/Plaything’ from Sound Characters (Making of the Third Ear), for instance, this is due to the use of gated arpeggios of pure tones. The DPOAE appears in the gaps between pure tone bursts, creating an intriguing dissonance between the acoustic sound ‘out there’ and the DPOAE sound ‘in here’, as the sounds segregate into two separate timbral streams. Due to the use of short, bright pure tones, ‘Headrhythm’ seems to reference the world of psychoacoustic demonstration CDs that come attached to academic textbooks. The DPOAEs are very much used as a disorientating effect rather than for its formal possibilities, and this is underscored by the sharp, abrupt shift into the very different sonic territory of ‘Plaything’. It feels as though the change in atmosphere is employed to somehow ‘justify’ the appearance of the preceding crude, possibly ‘unmusical’ arpeggios.\(^2\)

Jacob Kierkegaard, on the other hand, makes the DPOAE the explicit subject of an entire 38-minute work, Labyrinthitis. This approach bears the most formal similarity to Correlation Number One since the ear itself dictates the way the entire piece evolves. The piece begins with pairs of sinusoidal tones used to excite a DPOAE in the listener’s ears. The precise frequency of the DPOAE produced is then used to supply the next acoustic tone which will be introduced, meaning that the ear becomes an active participant in the work - it ‘tells’ the piece which frequency to introduce next and the work responds, creating a cascade of tones that mirrors the shape of the cochlear [11].

4. ‘VIRTUAL’ CONTROL

In the discussed pieces, the DPOAE is used in one of two ways, being either a) generated seemingly arbitrarily to create a disorientating effect (Amacher), or b) controlled, but only in pitch and intensity (Kierkegaard, Niblock). However, in my own experiments I soon discovered much more nuanced control was available just by manipulating the stimulus frequencies. In fact, under the right conditions the DPOAE can be treated just as if it were any signal, meaning that many of the ‘classic’ electronic music techniques are theoretically achievable. (The sound, after all, a physically observable fact that can be recorded with microphones.) Such experimentation required a system whereby one could apply processing techniques directly to the ‘virtual’ sound itself; the acoustic tones would be generated automatically within a range set by the user. This means that the composer would apply pitch, duration, and processing/timbre control to the absent tone that the listener’s ears would produce rather than the acoustic tones produced by the speakers (which would be generated mathematically as described below). The process thus became akin to ‘composing with absent tones’.

5. FORM THROUGH LIMITATION

Much of the literature on psychoacoustics treats the cubic difference tone, \(2f_1 - f_2\), as the most prominent otoacoustic emission. For musical purposes it proved to be unsuitable due to masking effects, and therefore the quadratic difference tone was used, which occurs at \(f_2 - f_1\) when \(f_1 < f_2\). Since ‘virtual’ control of the resulting frequency rather than the stimulus tones was needed, the calculation was inverted in max/MSP (see fig.1).

![Figure 1](maxMSP.png)

**Figure 1.** Sketch showing how each sinusoidal pair is generated in max/MSP.

The creative process from this low-level stage now took on an interesting character, whereby formal decisions were increasingly being imposed from ‘outside’ owing

---

\(^2\) Amacher herself expresses a similar sentiment in a conversation with Thurston Moore: [http://www.youtube.com/watch?v=SRrksHbVf8](http://www.youtube.com/watch?v=SRrksHbVf8)
to the capabilities of the hearing system and listening comfort. For example, we know that our ears would not respond in the same way irrespective of stimulus frequencies, since the cochlear is a frequency-specific organ [8]. A 100Hz DPOAE of f2 – f1 can be produced with countless combinations of frequencies, but the result depends upon intensity and how small the ratio between frequencies is. The smaller the ratio the better the resulting DPOAE will sound, but since this obviously results in combinations of higher frequencies, greater intensity is required for the effect to be audible. Thus, there is a logarithmic trade-off between frequency and amplitude which I found to be most pleasingly resolved around the area where the cochlear is most sensitive (2Khz – 5Khz), since the lesser amplitude is required for the effect to ‘work’ the more comfortable will it be to the listener. Even within this band, however, a SPL of approx 80dB was necessary for effective stream segregation to take place. This meant that the crucial question of the work’s duration became ‘how long could the piece be listened to, in comfort?’ An upper limit of 9 minutes was found to be most agreeable, since it allowed time for the process to reveal itself, stream segregation to take place and the form to be established before fatigue set in.

### 5.1. Stream segregation

**Correlation Number One** begins quietly, with single pure tones of between 3000Hz and 4000Hz fading in and out subliminally. Over the course of a minute or so these tones accumulate, separated by intervals prescribed by the input note (so a ‘G1’ at 3500Hz would prescribe a f2 frequency of 3695Hz); first pairs, then threes, fours and so on until the entire 8-channel array is producing a separate tone that reinforces the requested difference tone, or sequence of tones. When the fourth reinforcement tone is introduced, we at last begin to hear a melody being developed, and at this moment it becomes progressively harder to ‘hear’ the acoustic tones. Stream segregation by timbre and by space takes place, as we become aware of a stream that is both rougher in timbre and closer to us space, which is the DPOAE. Over the course of the piece, the effect becomes so pronounced that something like a loudspeaker-based equivalent to headphone listening is produced, whereby the DPOAE is localized distinctly around the head and the acoustic tones become like the noise outside, only partially audible. To aid segregation, light tremolo and vibrato was applied to the ‘virtual’ tone, which gives the DPOAE a timbral quality reminiscent of the sustained electric guitar tones that can be produced with an E-bow.

The form of **Correlation Number One** follows the technical method described above in the sense that it is not the acoustic tones that are the aesthetic object, per se, but the DPOAEs. The pitch of the DPOAE is derived from pairs of pure tones that, although not arbitrary, could feasibly be produced by other combinations within the range of 2-5Khz, whilst maintaining the essential quality of the piece. They act as stimulus tones, sound waves whose job it is to ‘play’ the ear. The relationship between the two elements might even be metaphorically linked to the relationship between performance and the score, in the sense that the stimulus tones are the immanent form of the work, whilst the ear ‘performs’ the piece, differently for each person.

### 5.2. Further work

In the application of tremelo and vibrato effects directly to the DPOAE, **Correlation Number One** touches on the possibility of a ‘DPOAE-synthesis’. This is a fruitful area for further exploration, even when implementing only ‘classic’ time-domain synthesis techniques. Synthesis of more complex sounds, however, requires more advanced techniques, and so the work being pursued now involves a more direct mathematical solution to harnessing the ear's nonlinearity.

### 6. CONCLUSION

The practice of composing with absent sound is predicated on a tension between ‘real’ and ‘virtual’ sound; a tension which, in **Correlation Number One**, spills over its bounds, becoming the focus of the entire piece via the segregation of ‘inner’ and ‘outer’ sound into two competing streams. Like the black circle we see if we stare at the sun too long, it is as if the pure tones are too bright and the hearing system ‘steps in’ to block our exposure to them, producing a negative effect in its place. It is tempting to say a shift from outer-sound (objective) to inner-sound (subjective) is made, but the hallucinatory quality that emerges serves merely to underscore the chimerical nature of sound experience in the universal. There is no absolute, fixed point, from which we can totally describe sound; it is always location-specific, produced in and by the subject and thus, forever incomplete. The auditory world is populated with temporal, intangible ghosts, whose ‘true’ identity is always betrayed by reflections and perceptual illusions. **Correlation Number One** magnifies the insurmountable screen before sound ‘as it is in the world’ and sound as it is ‘given’ to us by the auditory system. Against Cage, it highlights how there is no ‘sound in itself’, at least not one we can know about. ‘We cannot creep up on the object from behind’, as Hegel put it, since the only world available is that which is given to us by our perceptions, a partial representation [12]. Like in an impressionist painting where the scene represented is blurry and obscured and we are presented with the thrill of sensation only, **Correlation Number One** offers a ‘sensation of a sensation’; we hear hearing take place, as a further veil is placed over the real, and perception’s innate hallucinatory quality is magnified.

### 7. ACKNOWLEDGEMENTS

I wish to thank Gary Kendall for his assistance throughout the creation of this piece and for his comments and suggestions on the text.
8. REFERENCES


