

AN ANALYSIS OF JONATHAN HARVEY'S *SPEAKINGS* FOR ORCHESTRA AND ELECTRONICS

UN ANÁLISIS DE *SPEAKINGS* DE JONATHAN HARVEY PARA ORQUESTA Y ELECTRÓNICA

Gabriel José Bolaños Chamorro

Correo electrónico: gjbolano@asu.edu

Fecha recibido: 13 de noviembre de 2020

Fecha de aprobación: 12 de enero de 2021

* Gabriel José Bolaños composes solo, chamber and orchestral music, often with electronics. He holds a PhD in Music Theory and Composition from University of California-Davis and a BA in Music from Columbia University. He is Assistant Professor of Music Composition at Arizona State University, where he teaches courses in composition, acoustics and music technology.

Abstract

In recent years an increasing number of composers have used speech as source material for instrumental, electronic and electroacoustic music. This article examines this particular intersection of music and language through an analysis of Jonathan Harvey's *Speakings* for orchestra and electronics. I attempt to understand how Harvey made an orchestra sound like a human voice by analyzing his use of technology and his compositional techniques, particularly as they relate to existing theories of speech perception, acoustics and articulatory phonetics. This technical achievement is then placed in its broader musical context to examine the role that speech-sounds play in this piece, and the implications of hearing an orchestra speak in the context of this work's narrative.

Keywords: electroacoustic, orchestra, perception, music and language, phonetics.

Resumen

En años recientes, hubo un gran número de compositores que ha utilizado el habla como fuente para la composición de obras instrumentales, electrónicas y electroacústicas. Este artículo examina la particular intersección entre música y lenguaje a través del análisis de la obra *Speakings* de Jonathan Harvey para orquesta y electrónica. Analizando las técnicas compositivas de Harvey y el uso que hace de la tecnología, en particular como ambos se relacionan con teorías existentes de la percepción del habla, del acústica y de la fonética articulatoria, se intenta entender como el compositor consiguió que un sonido orquestal sonara como la voz humana. Este logro técnico es finalmente puesto en un contexto musical más amplio para examinar el rol que el sonido del habla juega en esta pieza y las implicaciones que conlleva escuchar una orquesta hablar en relación con la narrativa de este trabajo.

Palabras clave: electroacústica, orquesta, percepción, música y lenguaje, fonética.

1. Introduction

Jonathan Harvey's *Speakings* (2007-08), for orchestra, 11 amplified soloists and electronics, was commissioned by the BBC Scottish Symphony Orchestra and IRCAM with funds from Radio France. It is the third piece in a trilogy "referring to the Buddhist purification of body, mind and speech" (Harvey, 2008, Programme Note). *Speakings* is a realization of Harvey's long-time obsession with making an orchestra speak: "It's as if the orchestra is learning to speak, like a baby with its mother, or like first man, or like listening to a highly expressive language we don't understand" (Harvey, 2008, Programme Note). The BBC Scottish Symphony Orchestra, conducted by Ilan Volkov, premiered *Speakings* on 19 July, 2008 at the Royal Albert Hall in London, as part of the BBC Proms. Harvey collaborated closely with IRCAM technicians Gilbert Nouno and Arshia Cont on the computer music realization of this piece.

Speakings is rooted in Harvey's deeply personal and idiosyncratic spiritual beliefs that combine Buddhism, Christianity, and mysticism. The piece draws on many ideas that have been central to Harvey's compositional thinking since the 1980's: the use of bells, the timbre of the human voice, and a sophisticated use of live processing and computer-assisted-composition. This article only briefly touches on the technology used in this piece. For a detailed discussion of the electronics, please see Nouno et al. (2009).

The most unusual and perceptually salient feature of *Speakings* is hearing the instrumental ensemble produce anthropomorphic utterances. Screaming and cooing baby sounds, chatter, exclamations, and angelic mantras are just a few of the linguistic (or proto-linguistic) sounds present in this piece. Throughout the work, Harvey explores the timbral and perceptual distance between the human voice (heard primarily through live processing of the 11 amplified soloists) and the orchestra. The disembodiment and unintelligibility of the human voice, heard *through* an orchestra, and spatialized around the concert hall, creates extreme ambiguity that underpins the work's narrative.

The work is divided into three movements, played attacca: the first is "an incarnation... descent into human life" (Harvey, 2008, Programme Note) and is based largely on infant sounds, as if the orchestra is "learning to speak." The second, and longest, movement "is concerned with the frenetic character of human life in all its expressions of domination, assertion, fear, love etc." (Harvey, 2008, Programme Note). This movement concludes with a simple musical gesture modelled on a Buddhist mantra that represents "original, pure, speech. The OM-AH-HUM is said to be the womb of all speech" (Harvey, 2008, Programme Note). The final movement is a calm, largely monodic plainchant that concludes the work's dramatic arc. Harvey describes this as a "hymn which is close to Gregorian chant. There is often a single monodic line reverberated in a large acoustic space" (Harvey, 2008, Programme Note).

A formal large-scale process is clearly audible: 1) the emergence of voice-like gestures (learning to speak), 2) the development and growth of this material (domination,

assertion, fear, love, frenetic chatter, etc.), and 3) a final evolution away from speech towards the Buddhist mantra that finally evolves into plainchant. These three sections are not discrete states, but rather points along a spectrum: Harvey's gradual introduction of explicit voice-like material is accompanied by an even more gradual preparation of the mantra and plainchant through a bell-like sound orchestrated with piano, tubular bells and often trombone. While this narrative mimics language acquisition (learning to speak) and spiritual transcendence (the mantra and plainchant), Harvey constantly juxtaposes previous gestures with latter ones, creating a cyclicity that disrupts this directional process through continuous foreshadowing and recollection.

This analysis will begin with a discussion of what David Evan Jones calls "voice-likeness" (1987) to understand how Harvey used the timbral and rhythmic elements of speech to generate voice-like musical material. I will examine the evolution and transformation of these voice-like gestures, and how they relate to the Buddhist mantra and the plainchant, to further understand the work's narrative. Many of the compositional techniques used in this piece are rooted in so-called Spectral music, particularly Harvey's focus on timbre, his use of spectra to create timbre/harmony using instrumental additive synthesis, and his use of Computer-Assisted-Composition (Fineberg, 2000). The score to *Speakings* lacks measure numbers, so examples are referenced by page numbers or rehearsal letters. The adjective "solo" is used to distinguish between the concertino and orchestral instruments. The pitches notated in the MIDI keyboard are not sounding pitches: this keyboard is used to trigger and control the real-time electronics.

2. Voice-likeness as an analytical tool

Composer-theorist David Evan Jones defines voice-likeness as a sound or an aspect of a sound that might "...cue listeners' associations with the human voice in a given context" (1987, p. 140). Voice-likeness is closely related to source recognition: a listener's ability to recognize the source of a sound. Perception of voice-likeness depends both on the listener's mode of hearing (Bailey et al., 1977) as well as the acoustic properties of the sound. Jones argues that "...because a signal may be voice-like in one aspect (e.g. vibrato rate or intonation contour) and unvoice-like in another (e.g. timbre of the glottal source etc.) it is often more useful to refer to voice-like *aspects* of a sound rather than to voice-like *sounds*" (1987, p. 140).

Timbre is particularly important for speech perception: "When we describe the acoustic attributes necessary for a sound to be perceived phonetically, we are, in general, describing characteristics associated in music with timbre" (Jones, 1987, p. 147). Harvey leverages this important perceptual cue to create voice-likeness through real-time spectral processing of the 11 soloists using the visual programming language Max/MSP¹. This software imposes the spectral envelopes of various vocal

¹ Harvey used the *Gabor* library in Max/MSP, which uses Linear Predictive Coding to alter the spectrum of an input to resemble a target sound.

recordings onto the solo instruments during live performance, molding and shaping the instruments' timbres to resemble the complex timbral structures of speech (Schnell and Schwarz, 2005). Manipulating source-recognition is central to Harvey's approach to composing with real-time electronics: "With live electronics, when electronics are performed in realtime like instruments and combined with instruments (or, of course voices), two worlds are brought together in a theater of transformations. No-one listening knows exactly what is instrumental and what is electronic anymore" (Harvey, 1999b, p. 80).

The perceptual distance between voice-like and unvoice-like sounds is not linear nor continuous (House et al., 1962). One can easily create a linear acoustic continuum between voice-likeness and unvoice-likeness through simple interpolation (as Stockhausen did in *Gesang der Jünglinge*), but this will not be perceived as a smooth, continuous transition between speech and non-speech. Instead, listeners will perceive a sudden change from unvoice-likeness to voice-likeness (or vice-versa) when the sound passes a cusp of categorical perception. In short, the perceptual difference between speech and nonspeech is categorical: differences *within* each category are minimized while differences *between* categories are more pronounced. Voice-likeness can, therefore, be achieved (or explained) by understanding where this boundary lies, and the nature and number of acoustic cues necessary to cross this border.

Harvey also creates voice-likeness through the orchestration of formants in the orchestra. Formants are frequency peaks in the spectrum of the voice created by resonances in the vocal tract, which acts as a complex filter that alters the sound produced by the larynx. These resonances change with the motion of the tongue, lips and velum. Formants are independent of the fundamental frequency of the voice (see fig. 1). Each vowel is distinguished by multiple formants that vary by frequency and amplitude, but vowels can be accurately perceived with only the two lowest formants (Ladefoged, 2006 and 2010). Furthermore, formants emphasize certain frequency bandwidths, not discrete pitches, and can vary greatly depending on gender, age, body-type, accent and many other factors. For decades, linguists have been unable to find consistent invariants that directly correlate acoustic structures to phonemes, making it difficult to rely on spectral analysis to understand speech perception. This issue, known as the lack of invariance problem, means that a single phoneme may have different acoustic properties depending on the context, coarticulation, accent, speaker, emotion and prosody of the speech (Blumstein & Stevens, 1981, Samuel & Tartter, 1986, Stevens & Blumstein, 1979). As such, attempting to map the formants orchestrated throughout *Speakings* to exact vowels has proven impossible.

Rather than focus only on the frequencies of these formant gestures, I have found it useful to adopt a broader gestural analysis that combines texture, rhythm, range, dynamics and spectromorphology (Wishart, 1996 and Smalley, 1997) to better understand the orchestration and behavior of speech formants in the orchestra. In

general, orchestrated formants in *Speakings* are heterophonic and often rhythmically complex, emulating speech rhythms. Since formants are produced in the vocal tract by filtering a fundamental tone, they are much quieter than the fundamental, and tend to be much higher frequency (see fig. 1b). Mirroring the physiology of the vocal tract, Harvey tends to orchestrate formants in the strings and high woodwinds at low dynamics, and usually over a lower voice that represents the fundamental frequency of speech. The fundamental, or carrier, is also frequently manipulated with the live electronics to further shape its timbre to resemble speech.

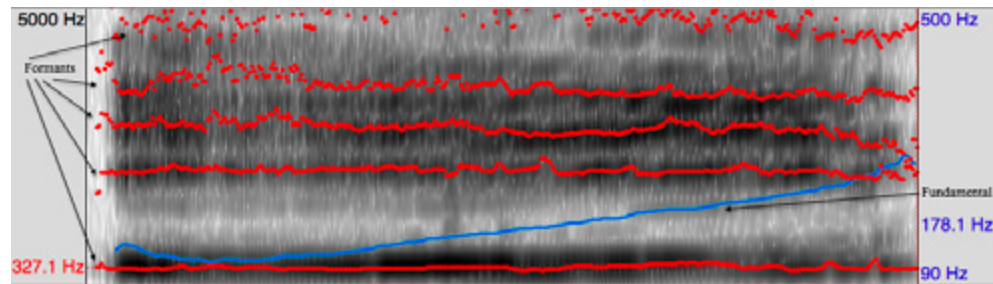


Fig. 1: a sonogram of a male singing the vowel [i] with an ascending glissando. This illustrates the independence between formants and fundamental frequency. The frequency range of the formants is on the left, the frequency range of the fundamental is on the right.



Fig. 1b: approximate frequency ranges of English vowel formants for men, women and children (Ladefoged 2010), expressed in standard notation.

3. Voice-Likeness in *Speakings*

By reframing speech as music, and creating timbral ambiguity that blends familiar orchestral instruments and the human voice, Harvey attunes listeners to focus on the beautiful, complex sounds of language rather than the meaning of speech: “In listening to ordinary speech, our attention is often occupied with the process of decoding the phonetic and syntactic message to the extent that we have little attentional capacity left to attend to the speech sound” (Jones, 1987, p. 143). This section will discuss various techniques Harvey used to achieve voice-likeness in order to better understand

the technical aspects of composing with language as music, and to comprehend how voice-likeness affects the work's narrative. Jones comments: "The acoustic difference between voice-like and unvoice-like is often very small; the psychological difference is often very great" (Jones, 1987, p. 141).

Incarnation and Learning to Speak:

The first explicit voice-like gesture in *Speakings* is the sound of a baby screaming (p. 7-8). The gesture begins with strings playing three simultaneous tonal harmonies: a C-major triad in the second violins (3 desks, *con sord*), an F# major triad in the first violins (3 desks, *con sord*), and a C-minor 7th chord in first inversion in the violas and in the violin, viola and cello soloists, all playing artificial harmonics tremolando (see fig. 2).



Fig. 2: three superimposed tonal harmonies preceding the first baby scream (p. 7).

The resulting harmony sounds kaleidoscopically triadic: the C-minor 7th chord is orchestrated most prominently, but is slightly shaded by the other two triads. The combination of C-minor 7th with the root of the F# major triad resembles a C-half-diminished 7th chord. The remaining pitches are all chromatic neighbor tones to this harmony (Gb, C# and Eb).

This harmony is both ambiguous and recognizable: listeners perceive the obvious timbre of the strings, and can hear the color of triadic harmonies despite being separated by a tritone. On the last beat of p. 7, the timbre of the baby scream is suddenly imposed on the string soloists (triggered by the MIDI keyboard), and orchestrated in the first violins (a very high chromatic cluster that mimics the second formant of the scream (p.7-8)). It is then spatialized around the concert hall, and immediately followed by electronic reverb as well as orchestrated reverb in the second violins, before fading away into the breathy alto flute motive that opened the piece.

This moment is the "incarnation" Harvey referenced in the program notes, the first pre-linguistic utterance heard through the orchestra. It also demonstrates many techniques that he uses throughout the work. Harvey frequently uses closed-position

triads as building blocks for more complex sonorities. While they have no tonal function, they create harmonic recognizability, and bridge the gap between timbral and triadic gestures. Similar triadic and tonal coloration is used extensively throughout the third movement to color the plainchant melody, when speech has been “...married to a music of unity” (Harvey, 2008, Programme Notes).

The first baby scream also illustrates the way Harvey orchestrates formants. By analyzing frequency and amplitude information of the speech sound, and assigning these frequencies and relative amplitudes to the instruments in the orchestra, Harvey re-creates the spectrum of the target sound through instrumental additive synthesis. In this example, the violins are orchestrated as clusters that mimic the nature of formants as frequency-bands rather than discrete pitches. The melodic contour of the cluster and final descending glissando also mimic the timbral morphology of the baby scream (fig. 3).

The image shows a handwritten musical score for Violins 1 and 2. The Violin 1 part (labeled 'Viol 1') consists of six staves (numbered 1-6) and a 'Soli 1' section. The Violin 2 part (labeled 'Viol 2') consists of two staves and a 'Soli 2' section. The notation is dense with clusters of notes, often marked with '3' or '8', and includes various dynamic markings such as 'fz', 'p', 'pp', and 'ppp'. Slurs and glissando lines are used to indicate the movement of the clusters. The score is written in a clear, legible hand.

Finally, Harvey also plays with ambiguity and source-recognition, in particular the sudden contrast between a highly recognizable sonority (strings playing a somewhat triadic harmony) and a completely new sound that moves throughout the concert hall and between the orchestra and concertino. The nature of this sound is unmistakably recognizable as an infant scream, but its source is made ambiguous by the electronics, spatialization and orchestration of formants. This ambiguity between acoustic and electronic sounds is conterminous with the ambiguity between music and speech that

is so important to this work's narrative. Harvey discusses his views of this interplay between acoustic and electronic sounds:

The art of live electronics highlights the unity of ambiguity still further. When electronics are performed in real time like instruments and combined with instruments or voices, the two worlds merge in a theater of transformations and legerdemain. No one listening knows exactly what is instrumental and what is electronic anymore. Legerdemain deceives the audience as in a magic show. When they lack connection to the familiar instrumental world, electronics can be overwhelmingly alien—other, inhuman, inadmissible, dismissible (like the notion of flying in a rational world). When electronics are seamlessly connected to the physical, solid instrumental world, however, an expansion of the admissible takes place and the “irrational” world is made to belong. If electronic sounds are completely separate from traditional instruments, they may as well be on the moon: there can be no measurement of interval and consequently no music (Harvey, 1999a, p. 62).

The development of pre-linguistic utterance (baby screams and babbling) into more complex voice-like patterns drives the first movement. As the shrieking disappears, the rhythms and contours of the baby babble become increasingly more complex. This process mimics speech acquisition in infants, and creates the sensation that the orchestra is indeed learning to speak, as Harvey described in the program notes.

Harvey uses the oboe and alto-flute soloists as carriers for the spectra of infant sounds through this movement. The low range of the oboe (played *mf*) is frequently used for the gentler babbling and cooing sounds, while the extreme upper register (played *ff*) lends itself well to screaming and shrieking. This particular passage (fig. 4) quickly alternates between the two:



Fig. 4: solo oboe alternating between baby babbling (low register, *mf*) and screaming sounds (extreme upper register, *ff*) (p. 10).

The glissandi between pitches imitate the natural speech contours of the human voice, a technique repeatedly used to create voice-likeness in the intonation of instrumental gestures. The oboe acts as the fundamental frequency of the baby's voice, imitating its contours and rhythms (prosody), while the timbre is altered in real-time by the electronics (phonetics). Harvey complements this again by orchestrating the formants of this gesture in the first violins.

Extreme registral changes in the oboe exploit source-recognition to help create voice-likeness. In the very high register, the oboe's timbre becomes almost unrecognizable, making listeners more amenable to perceiving the timbre of an infant scream instead of (or combined with) an orchestral instrument. By simultaneously moving away

from an easily recognizable instrumental timbre *towards* a vocal timbre (using live processing), Harvey successfully induces the listener to perceive vocalizations emanating from the concertino despite the obvious absence of a crying baby onstage.

The following is a list of all baby sounds in the first movement:

- p. 7: first baby scream, with resonance on p. 8
- p.9: second baby scream in the violins, followed by babble in the violas and celli.
- p. 10: third baby scream: the first time it appears in the oboe, followed by babble
- p. 11-12: baby scream reverb: both in electronics and orchestrated reverb
- p. 13: baby babbling in alto flute solo.
- p. 14: final baby scream, and babbling in oboe solo, accompanied by formants orchestrated in the strings.
- p. 17: baby crying. oboe is doubled in unison with violin.
- p. 18-21: increasingly more complex babbling in the oboe and alto flute soloists, with formants in the strings
- p. 21-23: the string soli are introduced as rhythms get more complex and dense.

The first movement ends with the string soloists playing voice-like motives together with the oboe and alto flute as the infant timbre fades away. Harvey does not use infant-sounds again until the end of the piece: some light babbling in p. 95 (rehearsal H), and the final gesture of the work is a very long, sustained, highly reverberant fade-out of infant cooing triggered by the MIDI keyboard (p. 103). The performance note for this gesture states “sustain ‘Baby Cooing’ to end... fading till ‘real’ baby sound totally dissolved into the orchestra” (p. 103) Concluding the piece with subtle baby cooing that dissolves into the orchestra reinforces the sense of unity in the final movement while also recalling the work’s long trajectory and narrative. The final infant sound helps listeners recall how far we have come since the incarnation, and imposes a cyclicity to the work’s linear narrative of speech acquisition followed by chaos and finally transcendence.

Infant sounds are an important motive throughout the piece: they are the most universally recognizable voice-like gestures, and are also the most emotionally communicative. The emotional prosody of an infant’s vocalizations, especially the contrast between screaming, crying, cooing and babbling, is understood by all listeners regardless of native language or culture. Harvey discusses pre-linguistic sounds as they relate to his faith:

Psycholinguistics, particularly in its movement from Lacan to Kristeva in Paris, finds deep significance in the psyches of preverbal children. They are, in this theoretical perspective (which sometimes seems amazingly close to the Lankavatara Sutra and other ancient Buddhist texts), a key to understanding the constitution of the subject, the exile from the garden of Eden. Our consciousness before and after we acquire language parallels human awareness before and after the Fall. Preverbal existence is empty, but full of richness (Harvey, 1999a, p. 49).

Short-Short-Long Rhythmic Motive

Harvey also uses motives that are voice-like in only certain aspects, which lie between speech and music on the voice-likeness spectrum. These are often recognizable instrumental timbres playing speech rhythms or contours. The most important of these is an ascending short-long rhythmic motive that resembles iambic speech rhythms. This motive is frequently at the head of longer voice-like melodies that use glissandi to mimic the contours of speech. The short-long motive is stated at the beginning of the piece:



Fig. 5: opening melody of the piece in the solo violin introduces the recurring short-long rhythmic motive (p. 1)

In the opening of the second movement, the motive is stated as a short-short-long rhythmic motive (fig. 6):



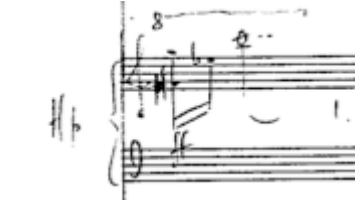








Fig. 6: short-short-long rhythmic motive in the solo oboe at the beginning of the second movement (p. 24).




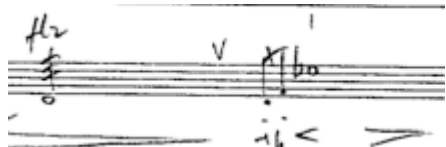

These closely related rhythmic cells are used prominently through the work. They often precede an upward leap of a 6th or 7th that slides down; also a common contour in speech prosody. The motive's simplicity and recognizability make it an ideal marker to highlight voice-like melodies it precedes: near the end of the second movement it is frequently used to mark a return to voice-likeness following sections that push closer and closer to the mantra. The result is an ebb and flow of gradual departure

and sudden return, with each departure reaching further away from voice-likeness towards the mantra and plainchant. Table 1 traces and discusses prominent examples of this motive throughout the work.

Table 1: prominent statements of short-long rhythmic motive in *Speakings*

<i>page no.</i>	<i>example with description</i>
1	 <p>Solo violin: the beginning of the piece. Glissandi mimic the contours of speech: either the third formant of a male voice or the second formant of a female or child voice.</p>
3-4	 <p>Solo violin: same as above, but slightly more developed.</p>
5	 <p>Harp, doubled by strings: the first statement of the short-short-long rhythmic cell.</p>
14	 <p>Solo oboe: alternating between babbling and crying ranges previously discussed.</p>

<p>24</p>	 <p>Solo oboe: the beginning of the second movement. The melodies are related to the baby babble of the first movement, but with larger intervals that mimic speech prosody.</p>
<p>25</p>	 <p>Solo oboe: a slight variation. More variations of this gesture follow.</p>
<p>32</p>	 <p>Solo oboe: coincides with the first bell-sound (see section four).</p>
<p>47</p>	 <p>Solo oboe, doubled in the strings: marks a return to voice-likeness following the first statement of the mantra and plainchant.</p>
<p>50</p>	 <p>Clarinet.</p>

<p>63- 64</p>	 <p>Many repeated overlapping statements in the winds</p>
<p>64</p>	 <p>Violin solo: prominent statement that leads to development of voice-like motives in the solo strings (pp. 64-69)</p>
<p>77</p>	 <p>Solo oboe: final statement right before the arrival of the mantra.</p>
<p>86</p>	 <p>Bass flute solo: beginning of the third movement. In this case, it emerges from low flutter-tongue similar to the gesture that opened the piece.</p>
<p>102- 103</p>	 <p>Bass flute solo: interval contraction</p>

This rhythmic motive is also augmented and distilled to its simplest form during the mantra (fig. 7), which is a simple repetition of two quarter notes and a half note. At this point the rhythmic identity has been separated from the upward leap and the voice-like melodies that so strongly characterized this motive. It is no longer used to prepare voice-like melodies, but is now isolated and reduced to an elementary form in the mantra that represents the “womb of all speech.” This mantra will be discussed in section five.

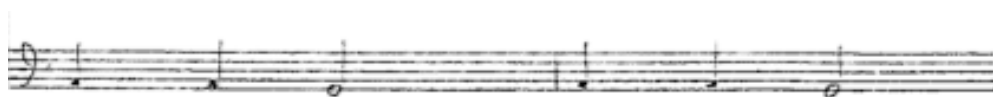


Fig. 7: short-short-long rhythmic motive in the mantra (pp. 78-84).

Frenetic Chatter of Everyday Life

At the beginning of the second movement (pp. 24-32), Harvey alternates between these voice-like melodies in the solo oboe and dense blocks of sound in the concertino that are heavily processed to sound like speech. Figure 8 clearly illustrates this alternation.

The dense blocks of processed sound in the concertino are the first instance of adult vocal-sounds in the piece, and these return throughout the second movement (p. 24, 25, 26, 27, 29, 30, 39, 47, 51, 53 and 54).² The dense clusters in the piano create rich overtones which are particularly well suited for being molded to speech sounds: the Gabor objects Harvey uses only emphasize or attenuate existing overtones in the carrier sound to mimic speech. A sound with a very rich, complex spectrum is more likely to have common overtones with the target speech sound than one with fewer overtones, and will therefore more closely resemble speech.

Fig. 8: the opening of the second movement: alternation between voice-like melodies in the solo oboe and blocks of processed sound in the concertino (p. 24).

² Some blocks are exact repetitions: the block on p. 24 repeats on p. 47, p. 25 repeats on p. 51, two blocks on p. 26 repeats on p. 53 and p. 54, and p. 30 repeats on p. 39.

This alternation between the solo oboe melody and concertino emphasizes different acoustic components of speech: the solo oboe employs nimble speech rhythms and contours to create voice-likeness while the blocks of sound are voice-like only in timbre. This alternation continues until the first bell-sound (p. 32), at which point these two techniques begin to fuse: complex speech rhythms and melodies are colored homophonically and heterophonically by larger subsets of the concertino and orchestra while being processed with the electronics. Furthermore, at this point the trombone begins to replace the oboe as the primary carrier of speech timbres (fig. 9). Like the oboe, the trombone can cut through dense orchestral textures, but its range is better suited to emulate adult male voices. It is also capable of producing complex voice-like glissandi, and its rich overtones are ideal for live processing with the Gabor library.



Fig. 9: the solo trombone takes over as the primary carrier of speech timbres (p. 35).

The most complex voice-like gestures in the piece occur between p. 48 and p. 69, after the work's first large structural arrival which foreshadows the mantra and plainchant (p. 46). This highly virtuosic section develops the "frenetic chatter of human life," and features extensive electronic processing of the concertino combined with orchestrated formants. Harvey writes in the program note that the frenetic chatter expands on his work *Sprechgesang*, composed just before *Speakings*.³

Harvey creates the sensation of frenzied, agitated speech by superimposing many relatively simple techniques. Imitation is used to create a sense of echo and reverberation: on the final measure of p. 51 the piccolo imitates the cello solo's theme from the previous measure 5 octaves higher with slight rhythmic variation. This imitation continues between the piccolo and the solo trombone on pp. 52-53. There is also extensive imitation in the winds on p. 64 following the large climax on the previous page, creating a blurred, unfocused dissipation of energy using the short-short-long rhythmic motive.

From pp. 48 - 63, Harvey orchestrates formants to create dense, complex and highly colorful heterophonic textures that explore the rapidly changing and nuanced timbres and rhythms of speech. These textures complement and expand on the solo trombone, which serves as the fundamental frequency and carrier of the voice-like sound, and is heavily processed by real-time electronics to resemble a male voice.

³ There are many parts of the second movement borrowed directly from *Sprechgesang*, including a shortened version of the oboe solo alternating with blocks of sound that opens the movement.

Conceptualizing the orchestra as a single, highly expressive vocal tract allows us to analyze these gestures through the lens of articulatory phonetics.⁴ The human vocal tract is the most complex and nuanced of all musical instruments, and employing its physiology as a model for controlling musical gesture has fascinating compositional implications. In-depth analysis of one excerpt from p. 48 (fig. 10) will demonstrate the techniques that Harvey uses to craft these formant textures throughout this entire section, and how they relate to the acoustics and articulation of speech.

The solo trombone is lightly colored by a solo cello that plays in unison with small chromatic and microtonal variations, and a solo viola at the octave (the first overtone). Low formants are orchestrated heterophonically in the first violins: as usual, they follow the melodic and dynamic contours of the trombone. They begin as clusters that mimic the frequency bands of formants, but on the second eighth-note of the triplet on the third beat, this cluster suddenly expands to occupy a wide range of the trombone's overtones (overtones 3 to 8, which also outline major triads above the fundamental and are, in this case, orchestrated as parallel major triads). For example, on the downbeat of measure 2 of this excerpt, the trombone and cello play a low G, the viola plays the first overtone, and the first violins play divisi overtones 2 to 5. During longer, accented notes such as this one, the formants coincide exactly with the trombone's overtones. On shorter and/or non-accented notes, the formants often diverge slightly from the trombone's overtones. This timbral phenomenon can be clearly observed in sonograms of speech: accented and longer vowels tend to have clearer, more stable formants, while the formants of shorter and less accented vowels are more susceptible to change either by assimilation or coarticulation of a nearby vowel or consonant (the process by which one sound becomes like a nearby sound).

Harvey also orchestrates fricatives in this passage (marked with circles in figure 10) in the orchestral percussion (sand block, cabaca, guero) and solo percussion (hi-hat, woodblock). These instruments all produce high frequency-bands of noise similar to fricatives such as "s," "sh," etc. The final two fricatives in this example (m. 2 and m. 4 on p. 48) coincide with rests in the orchestra and concertino. This delicate and subtle hocket between pitched and unpitched instruments requires careful and precise rhythmic control, and is an example of an unvoiced fricative such as "s" (as opposed to a voiced fricative like "z") eliding into a voiced vowel. The high formants orchestrated in the second violins act as a bridge between the low formants (vowels) and these orchestrated fricatives. This is another example of linguistic assimilation or coarticulation: if one says, for instance, the word "as" while closely observing the motion of the tongue, it becomes clear that the transition from the open vowel "a" to the voiced fricative "s" is smooth and continuous. The tongue gradually moves up to the alveolar ridge to articulate the consonant "z," as it is physically impossible for the tongue to instantly appear in this position without occupying the space in between the bottom of the mouth ("a") and the top ("z"). This motion results in a very rapid acoustic modulation between adjacent phonemes: the formants do not suddenly and discretely

4 For a good overview of articulatory phonetics, see Ladefoged 2006 and 2010.

Harvey also relies on question-answer phrasing to create the sensation of frenetic chatter. This is used throughout this section (pp. 48 - 69) to create the impression that instruments are speaking to each other: there is even a conversation between the trombone solo (with orchestrated formants and processing) and electronic voice-like sounds triggered by the MIDI keyboard (pp. 55-59). In a live performance, hearing these prominent electronic sounds emerge while very few members of the orchestra are playing (p. 59) creates an interesting visual dissonance that further obscures the origin of these voice-like gestures.



Fig. 11: one of the most unusual voice-like sounds in *Speakings* is acousmatic: a buzzing electronic timbre seems to speak (p. 71).

The buzzing, electronic timbre is very distant from the predominantly orchestral and electroacoustic palette of the work up to this point, but it maintains similarity through rhythmic and phonetic voice-likeness. Jones' observation about voice-like *aspects* of a sound explains how Harvey can create a sound that is simultaneously voice-like and unvoice-like. The surprising timbre is eventually blended with the orchestral timbral palette in p. 71 through doublings in the trombones (with cup and wah-wah mutes), and on the final measure of p. 72, with a similar triangle-wave-like timbre in the trumpets with straight mutes. By gradually blending this unusual timbre with the orchestra, Harvey reigns in the surprising outlandishness of this timbre and helps integrate it with the orchestra, effectively exploding the realm of acoustic possibilities at this structurally important moment of the piece. The statement on pp. 74-75 is the final explicitly voice-like gesture that uses electronic processing in the piece, and occurs immediately before the last statement of the short-short-long motive (p. 77 in the oboe) and the arrival of the mantra.

4. Bell-sounds

Bells and their rich, inharmonic spectra have been a source of inspiration for many composers including Claude Debussy, Jean-Claude Risset, Gérard Grisey and Tristan Murail. Harvey used bell-sounds prominently in *Mortuos Plango*, *Vivos Voco* (1980) and *Ritual Melodies* (1990). In *Ritual Melodies*, an acousmatic piece developed at IRCAM, Harvey used synthesized bells to help delineate form and mark interpolations between various synthesized instrumental timbres. *Mortuos Plango* was a very personal piece for Harvey, and made use of his son's voice together and the tenor bell at Winchester

Cathedral. By synthesizing and transforming these two sounds, Harvey was able to contrast the “dead voice of the bell” with the “living voice of the boy” (Harvey, 1981, p. 22). Like *Speakings*, this work employs bell overtones, spatialization, reverb, linguistic analysis, phonetics, triadic harmonies and golden ratio proportions to inform and control the interactions between and blending of voice-like and bell-like sounds. Bell-sounds play a more subtle yet structurally important role in *Speakings*.

Bell-sounds orchestrated in the piano and tubular bells act as a catalyst for change: they usher the transition away from voice-like frenetic chatter towards the calm, resonant mantra that concludes the second movement and represents the “womb of all speech.” Figure 12 is a reduction of the bell-sounds that lead up to p. 77, the structural turning point where speech has been exhausted and the mantra begins to emerge. The mantra, a slow repetitive sway between A and G, is an orchestration of the syllables “om-ah-hum,” (Cochard, 2008) and begins on p. 78.

The Emergence of Bell-Sounds

The first bell-sound is on the last measure of p. 32: a low major 2nd in the piano (C, D with pedal) doubled by contrabassoon, trombone, tuba, bass and tam-tam, and accompanied by the short-short-long rhythmic cell in the oboe. This orchestration is very low, rich, static, and resonant, especially compared to the high, rhythmic and almost agitated voice-like material that precedes it. This contrast firmly segregates these two theme-groups, and creates the impression that they are developing simultaneously yet on two separate planes.

While this is the first explicit bell-sound, these can be traced back to the very opening of the piece. The subliminal 17 Hz tone triggered by the MIDI keyboard on p. 1 is a sub-audible C# (although its overtones are indeed audible). This figure repeats on p. 3, p. 5 and p. 11, each time doubled by the bass. These are the lowest pitches in the first movement, and like the bell-sounds, they are low, sustained and spatialized notes surrounded by higher rhythmic voice-like material.

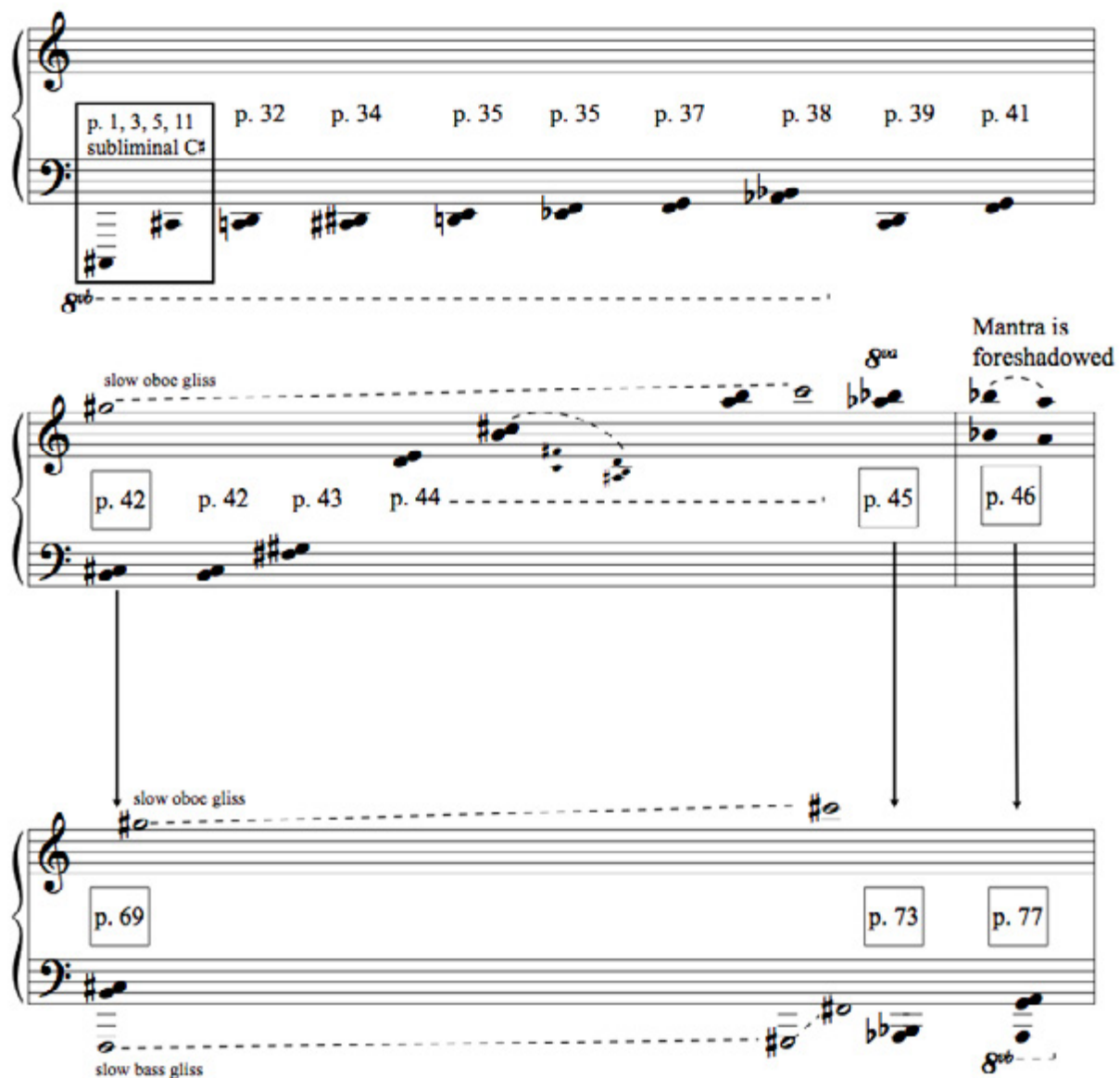


Fig. 12. reduction of bell-sounds (black noteheads). White noteheads are oboe and bass glissandi. Vertical arrows show formal parallelism between sections.

Furthermore, the ascent of the bell-like sounds from the C, D dyad on p. 32 to the high Bb, Ab dyad on p. 45 is exponential: it begins slowly and exponentially increases as it approaches p. 45 (see figure 12). The C# in the first movement is an extension of this exponential curve, a low asymptote from which the entire ascent emerges. Figure 13 shows the final statement of the subliminal C# on p. 11, accompanied by a half-step ascending glissando in the bass: a small gesture that illustrates the point of departure from this asymptote.

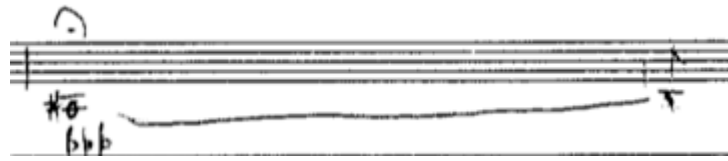


Fig. 13. Sub-audible C# tone, doubled by bass that glisses up (p.11)

The 17 Hz tones in the first movement are subliminal ancestors to the bell-tones. This C# is the sub-audible, obscured starting point of the bell-sounds' ascent (fig. 12) that eventually leads to the work's climax. In this respect, p. 37 is not the first bell-tone, but rather the point of emergence of a deeper, primordial, almost subconscious gesture.

Form, Parallelism and the Golden Mean

These bell-like sounds are important structural signposts through the second movement, and outline a strict formal parallelism between two large sections: pp. 42-46 and pp. 69-77. Each of these bell-sounds relates directly to a corresponding bell-sound in the other section: p.42 corresponds to p.69, p.45 to p.73 and p.46 to p.77. This relationship is marked with vertical arrows in figure 12, and is further outlined in the following chart:

<p>p. 42: (B, C#) followed by a diminished triad and a very slow ascending glissando in the oboe.</p>	<p>p. 69: (B-C#) followed by a diminished triad and a slow ascending glissando in the oboe and bass.</p> <p>Followed by a trombone melody with slow glissando</p>
<p>p. 45: (A\flat, B\flat) very high bell-sound. Coincides with end of oboe gliss.</p>	<p>p. 73: (A\flat, B\flat). Low bell-sound, also coincides with the end of the oboe gliss.</p>
<p>p. 46: (B\flat → A\flat) resolution down to A\flat, first statement of mantra in the electronics (English horn).</p> <p>Approximately coincides with the golden mean of the golden mean.</p>	<p>p. 77: (G, A). Followed by a trombone melody with slow gliss.</p> <p>Approximately coincides with the golden mean of the piece.</p> <p>Re-stated on p. 78 (G, A) The mantra, a slow sway between A and G, dominates the rest of the movement.</p>

This parallelism reflects an abstraction of the short-short long rhythmic cell on a macro level: like the final note of the rhythmic motive, the final bell-sound of each of these groups is structurally accented:

(B, C#)	(Bb, Ab)	(A → Bb)	
[p. 42 (high oboe gliss)	p. 45	p. 46]	→ Mantra, return to voice-likeness
[p. 69 (oboe + bass gliss)	p. 73	p. 77]	→ Mantra, transition to 3rd mvmt
(B, C#)	(Bb, Ab)	(G, A)	(A → G repetitive sway)

Furthermore, the arrival on p. 77 coincides approximately with the work's golden mean, and the arrival on p.46 is the approximate golden mean of this golden mean. The piece is 28 minutes long, so the golden mean is at approx 17:30 (28 min x .618 = 17:30). The golden mean of this duration is at approximately 10:42 (17:30 x .618 = 10:42) (see figure 14). The bell-sound arrival on p. 46 occurs at 10:25, and the arrival on p. 77 occurs at 17:00. While these do not coincide perfectly with the calculated golden ratios, they are very close. These numbers are based on timing of events in the only available recording of the piece, not calculated durations from the score.

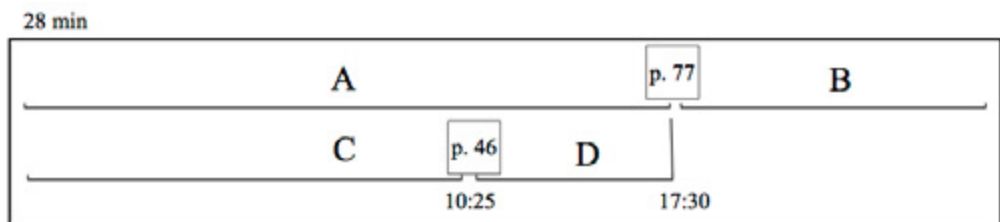


Fig. 14. approximate golden mean proportions articulated by two important structural arrivals on p. 46 and p. 77.

Proportionally, A is to B what C is to D, but these sections have different structural purposes. A is dominated by the development of voice-like themes and the emergence of bell-sounds. The arrival on p. 77 marks the end of voice-like gestures and the beginning of the mantra that eventually leads to the plainchant of the third movement (B). Within A, the arrival on p.46 briefly foreshadows the mantra and plainchant before returning to a thorough development of voice-likeness in D.

Harvey makes no mention of the golden section in his writings and interviews about this piece, but he has clearly articulated this proportional self-symmetry with the two aforementioned bell-like sections (pp. 42-46 and pp. 69-77). This is likely intentional, especially considering his use of the golden ratio in *Mortuos Plango Vivos Voco*, which is famously based on the human voice and the bell at Winchester Cathedral.

The first section of this parallelism (pp. 42-46) begins with a bell-sound (B, C#) and a long resonance orchestrated in the horns, strings and tubular bells. This resonance loosely outlines a diminished triad: the G and B \flat in the tubular bells combines with the C# (enharmonically a D \flat) in the horn to create an ambiguous diminished triad, similar to the one discussed in the previous section. A high G-# in the oboe slowly glisses up to a C# in p. 45, linking these two bell-sounds.

The second section (p. 69-77) begins almost identical to the first: the same bell-sound is followed by a similar resonance (also with a G and B \flat in the tubular bells outlining a diminished triad). The slow ascending oboe glissando is now accompanied by a low bass glissando (that ascends exponentially), creating a similar resonant space between the two bell-sounds on p. 69 and p. 73.

The two sections are framed by these pairs of bell-sounds with diminished-7th resonance and long, slow ascending glissandi (pp. 42-45 is similar to pp. 69-73). By employing clearly, almost obviously audible similarities, the differences between these sections becomes even more apparent: pp.42-45 is dominated by various ascending bell-sounds. This section is used primarily to develop this motive, and constitutes the exponential portion of the bell's ascent previously discussed. The corresponding section on pp. 69-73 does not develop the bell-sounds. Rather than an exponential ascent to A \flat , B \flat on p. 45, there is a sudden descent to the very low A \flat , B \flat bell-sound on p. 73. This section contains important developments of voice-like motives. The bell-sound on p. 69, along with its direct repetition on the second measure of p. 70, mark the first statements of the glissando trombone gesture that eventually lead to the mantra and plainchant (figure 15). This section also contains the buzz-speaking previously discussed; the most unusual example of voice-likeness in the work.



Fig. 15. two statements of the glissando trombone gesture following the bell-sounds on p. 69 (top) and p. 70 (bottom)

The most important differences between these two sections, however, are in the final, structurally accented bell-sounds that lie at the golden means (see figure 12 and figure 14). The first section (pp. 42-46) arrives at a very high B \flat , A \flat bell-sound, followed by a melodic resolution from B \flat to A \flat : this major-second foreshadows the swaying mantra that ends the second movement, and is a melodic expansion of the bell-sound's major-second harmonic interval. Furthermore, this arrival also foreshadows the plainchant of the third movement: in the third measure of p. 46, a pre-recorded English horn theme is triggered by the MIDI keyboard (figure 16). This motive is intervallically similar to the plainchant themes in the third movement, and the English horn is not used again until the final measure of p. 89, when it is featured prominently in the plainchant. By briefly foreshadowing both the mantra and the preceding plainchant at this local climax, and then returning to a prolonged section of frenetic chatter (p. 47-69), Harvey creates a non-linear narrative between the development of voice-like

themes and the emergence of the mantra and plainchant from bell-sounds. In short, this creates a sense of formal ebb-and-flow between these two contrasting ideas.



Fig. 16. English horn playback foreshadows plainchant of the third movement at first major structural arrival (p.46)

The second section (pp. 69-77), by contrast, concludes with the lowest orchestrated bell-sound in the work (using the lowest note on the piano). Excluding the sub-audible C# in the first movement, this is the lowest pitch in the piece, and recalls the original very low register and timbre of the bell sounds. This moment is also preceded by recalling the short-short-long rhythmic motive in the solo oboe, by now very familiar to listeners:



Fig. 17. short-short-long motive before large structural arrival on p.77

This statement of the motive is highly exposed, and played with heavy electronic reverb, imbuing it with an almost mystical sense of depth. This creates the sensation that the oboe and bell-sound are now occupying the same acoustic space. The result is a sense of resolution based on familiarity and unity: by recalling the familiar motive in the oboe just before the bell-sound in its original register, and physically unifying these two gestures through a common reverb, Harvey forges an important point of conversion and resolution between these two theme-groups.

Resonance and reverb are important features of these bell-sounds. They are always played with the piano pedal down, emphasizing the rich inharmonicity of clashing overtones separated by a major 2nd. The fundamental pitches are usually doubled in the orchestra, and as the bell-sounds develop they are followed by orchestrated reverb or echo in the tubular bells that becomes increasingly richer and more extended. These three examples (figures 18, 19 and 20) illustrate how Harvey gradually develops and expands the resonance, making it a larger and richer acoustic space as the bell-sounds develop:

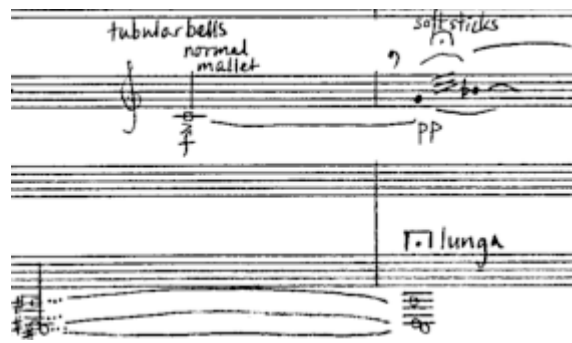


Fig. 18. bell-sound in the piano and resonance in the tubular bells (p. 34)



Fig. 19. bell-sound in the piano and more complex resonance in the tubular bells and piano (p. 42)

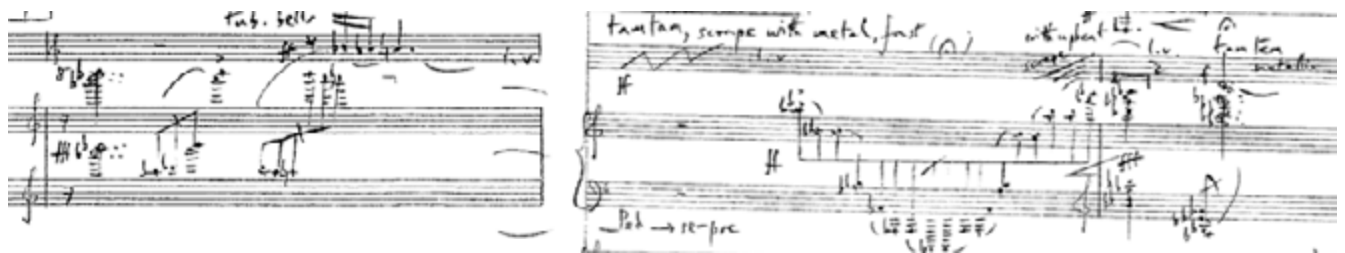


Fig. 20. bell-sounds in the piano and even more orchestrated resonance in the percussion and piano (p. 45 and 46)

This rich, reverberant orchestration also contributes towards the horizontalization of vertical sonorities through the prolongation and arpeggiation of harmonies during the orchestrated resonance that follow the bell-sounds. The increasing length and complexity of the orchestrated resonance creates the sensation of zooming into the resonance, pushing listeners deeper and deeper into the bell-sound, and transforming verticalities into melody. This technique also contributes to the seamlessness with which the mantra emerges from these bell-sounds in p. 46 and pp. 77-78: the mantra is simply a repeated descending major-2nd, the same interval that comprised every

preceding bell-sound and resonance. Orchestrated and electronic reverb contribute to the sensation that the mantra emerges from deep within the bell-sound, and further unifies these two highly contrasting musical gestures.

5. Mantra and Plainchant

The mantra marks the transition between speech (the first two movements) and song (the final movement) in order to unify these two contrasting sonic worlds. As Harvey writes, “Unification needs contrast with which to articulate itself” (1999a, p. 67). The plainchant in the third movement consolidates this unification by eliminating the “division of line against line, or music against listener” (Harvey, 2008, Programme Notes). The extreme contrast between frenetic chatter and the mantra and plainchant creates an even more powerful sense of resolution: “The greater the conflicts it successfully unifies, the more spiritual the music” (Harvey, 1999a, p. 52).

Mantra

In Buddhist mythology from India there is a notion of original, pure speech, in the form of mantras - half song, half speech. The OM-AH-HUM is said to be the womb of all speech (Harvey, 2008 Programme Notes).

The second movement concludes with an orchestration of the mantra “om-ah-hum” that is processed by the electronics. This mantra is calm, simple, repetitive and peaceful; an extreme contrast to the baby sounds and frenetic chatter of the first two movements. It is a simple ostinato that begins in the solo cello (p. 78) and gradually spreads through the entire orchestra and concertino until the end of the second movement (pp. 78 - 85). Figure 21 shows a harmonic reduction of this process. The harmonic expansion consists of a gradual unfolding up the harmonic spectra of A1 and G1, the two pitches that comprise the mantra⁵, which are also articulated by repetitions of the bell-sound in the piano and tubular bells (displayed in red). The result is a peaceful, meditative sway between A and G that envelops the listener as it grows out of the bell-sound and expands through the orchestra and the concert hall.

5 The mantra is played primarily an octave above these pitches, A2 and G2. The bell-sounds, however, are A1 and G1.

15 p. 81, rehearsal JJ

15 p. 82

15 p. 83

15 p. 84

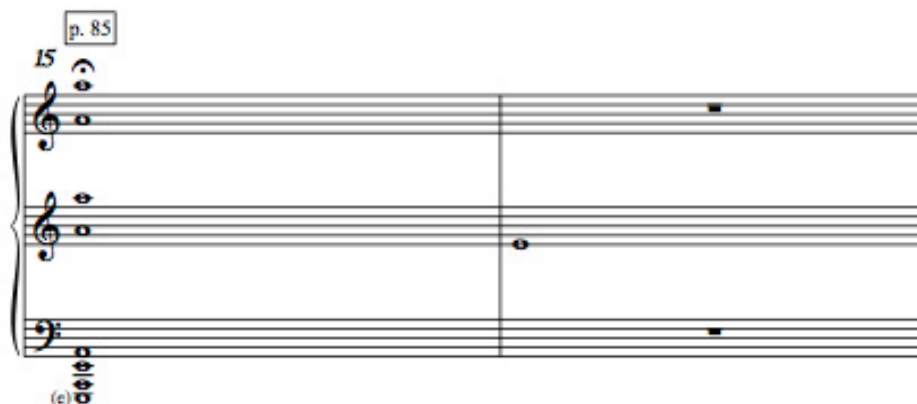


Fig. 21: harmonic reduction of the mantra om-ah-hum (pp. 81-85) that concludes the second movement. Bell-sounds orchestrated in tubular bells and piano are in red.

Harvey's views on spectralism, language and time have fascinating implications for understanding how the mantra fits into the work's narrative:

Spectralism, like harmony, is in essence outside the world of linear time. In music, time is articulated by rhythm; in psychology, time is articulated by the process of chopping up and arranging experience into language, which separates us from the primary world and joins us to the linear symbolic order (1999a, p. 40).

For Harvey, linear time as articulated by linguistic categorization is analogous to rhythm in music. By eliminating the linearity of rhythm in the mantra through extreme repetition of a simple rhythmic cell, Harvey is also undermining this corresponding linguistic categorization, and returning the listener to the "primary world" where "linear symbolic order" does not exist, and speech becomes pure sound. While this is not audible in the music, this interesting quotation offers a glimpse under the hood, into Harvey's personal opinions on the relationships between time, rhythm and language that may have informed his compositional thinking.

Plainchant and the Third Movement

Three major theme-groups dominate the third movement: plainchant, bell-sounds and breathy bass flute melodies with slow glissandi. The bell-sounds in this movement are less resonant and higher pitched than in the second movement: they are often

doubled in the celesta and harp, and are constructed from superimposed triads rather than minor-seconds (figure 22). The breathy bass flute motive (figure 23) opens this movement (p. 86), and re-articulates the middle section of the movement (pp. 94-95). The plainchant (pp. 89 - 100) is a single melody with harmonic and timbral color added: “throughout the plainchant passage... the bottom line is dominant and quite expressive, whereas the chordal coloration notes accompanying it above are very soft and without much dynamic range” (performance indication in the score, p. 90).



Fig. 22: bell-sounds in the final movement (p. 87).



Fig. 23: bass flute motive that opens the third movement (p. 86)



Fig. 23: bass flute motive in the middle of the third movement (pp. 94-95).

The third movement can be roughly divided into three sections: the first (pp. 86-94) opens with the bass flute motive before proceeding to bell-sounds and a few simple statements of the plainchant. The second (pp. 94-98) re-states the bass flute motive before proceeding to a denser, more developed plainchant punctuated by bell-sounds. In the final section (pp. 98-103), the plainchant ascends to the highest register of the orchestra, eventually fading into a very high A# (p. 102, rehearsal N). This final section contains no bell-sounds, and the plainchant is punctuated by voice-like gestures in the solo oboe (p. 98), solo clarinet (p. 99), solo piccolo (pp. 99-100) and finally the bass flute (pp. 102-103).

As a side-note, there are also unusual sweeping string gestures in the concertino throughout this movement. They are played *pianissimo* con sord, and instructed to maintain a tempo “completely independent of the conductor” (performance indication in the score, p. 89). Furthermore, they are instructed to repeat a long 9-measure phrase as often as necessary through pp. 89-94, also independent of the conductor. Later in the piece (pp. 96-98), the string soloists repeat gestures in boxes aleatorically, freely changing the order of the boxes. This independence between the concertino and the orchestra contrasts greatly with the complex, interconnected textures in the first two movements, and creates an interesting aleatoric, almost isorhythmic background for the plainchant melodies. These gestures are not prominent, and function largely as background and connective tissue between the three themes. These recall the maraca background in Tristan Murail’s *Ethers*: they are constantly present in the background such that their sudden absence creates a very strong accent. In this case, when the strings fade out on p. 94, there is a local structural accent to mark the return of the bass flute motive in the middle section of this movement.

The actual plainchant melodies are rhythmically and melodically very simple and calm, and they are “sung” either monophonically or homophonically by large subsets of the orchestra. Figure 24 shows a melodic reduction of the plainchant melodies. Intervally, they are constructed primarily with seconds and thirds, and occasionally fourths and fifths. These melodies often briefly and indirectly hint towards certain modal or tonal collections. Like the short-short-long rhythmic motive, they always begin with an ascending interval (usually a leap) followed by descending motion, a common contour of western vocal music. Most of the melodies consist of two periods with this general contour (denoted by dotted slurs in figure 24). Rhythmically, these melodies use predominantly eighth and quarter tones, with occasional dotted eighth or dotted quarter rhythms. These features create a strong continuity and consistency through the third movement, and contribute towards the calmness and stasis of these motives.

p. 89

p. 91

p. 92

p. 93

p. 95

p. 96 (rehearsal I)

p. 96 (rehearsal J)

Fig. 24. Melodic reduction of the plainchant (pp. 89-100).

At first the plainchant is simply doubled in unison (p. 89), but becomes increasingly more harmonically and timbrally colorful as it develops. There are many instances where modal or tonal collections are used non-functionally to add color similar to the kaleidoscopic triads previously discussed in section three. For example, on the second beat of the second measure on p. 90, Harvey orchestrates the pitches of G-phrygian (G A \flat B \flat C D E \flat F) in the strings above the plainchant melody to add an ambiguous modal shading to the end of this phrase. B-mixolydian and G-major are similarly orchestrated in the strings on the final measure of p. 92, A-mixolydian on the penultimate measure of p. 93, and B \flat -major is used in the final measure of p. 96. This simultaneous orchestration of entire modal pitch collections serves as another layer of orchestrational reverb: by merging all the pitches of familiar modes into a simultaneous sonority, Harvey is temporally uniting them into a single timbre/color, as if he is orchestrating only the resonance of these pitch collections.

As the plainchant motives ascend to the highest register of the orchestra, fragments of voice-like motives are recalled in the English horn (p. 98 and p. 99) and the piccolo (p. 99 and p. 100). A final, exhausted breath of orchestrated formants in the solo strings and high woodwinds (p. 101) precedes the final breathy bass flute motive (fig. 25). This gesture sounds like a shy, intervallically compressed synthesis of the plainchant, the first two movement's voice-like melodies, and the breathy flute glissandi that open the piece and the third movement.

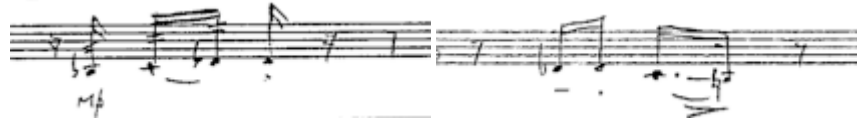


Fig. 25. final motive, in the bass flute (pp. 102-103)

Conclusion

This analysis was driven by a personal desire to better understand how language and speech can inform music composition. Harvey's *Speakings* is one exceptional example in a recent trend of works that explore this subject, and I hope this analysis may help others analyze and/or compose works that use speech or language in abstract or unusual ways. As I dug deeper into *Speakings*, I found myself repeatedly asking two questions: *how* and *why*? The first question, while technically challenging, is the most straightforward of the two. How did Harvey achieve extreme voice-likeness?

I examined the technical, technological and compositional means by which Harvey achieved voice-likeness by using an analytical framework that combines theories of speech perception with David Evan Jones' concept of voice-likeness and source-recognition. Technology, in particular computer-assisted-composition and live processing, was central to achieving this goal. Indeed, Harvey's unique technological needs led to the development of new computer-music tools at IRCAM specifically designed for the realization of *Speakings*. However, at its core this is fundamentally a question of forging resemblance: how to make one thing *seem* like another. Harvey achieved this through perceptual ambiguity of source recognition: unusual instrumental timbres such as extended techniques, unique orchestrations, or extreme registers allow listeners to dissociate the *sound itself* from the sound's physical *source* (the instrument). This permits greater suggestiveness in perceiving a sound as voice-like rather than, for example, oboe-like. Real-time processing then pushes the sound beyond the perceptual border separating speech and nonspeech, creating even more explicit voice-likeness despite the obvious absence of a human speaking onstage, or any recognizable words. In short, there is a simultaneous timbral shift *away* from an ordinary, familiar instrumental sound *towards* a voice-like sound that fundamentally calls into question the nature and origin of the sound. This ambiguity is further emphasized through spatialization and reverb that obscure the physical origin of the sound.

I am convinced that this perceptual ambiguity is *why* Harvey chose to explore the use of language as music in this highly personal and spiritual piece that fuses seemingly dialectical pairs like music and language, Buddhist mantra and plainchant, uncertainty and clarity, concertino and orchestra, and electronics and acoustic instruments. This work is one more elegant manifestation of Harvey's obsession with synthesizing seemingly opposing forces through ambiguity.

References

Bailey, P., Summerfield, Q., & Dorman, M. (1977) On the Identification of Sine-Wave Analogues of Certain Speech Sounds. *Haskins Laboratories Status Report on Speech Research* SR-51/52.

Blumstein, S. E., & Stevens, K. N. (1981). The search for invariant acoustic correlates of phonetic features. *Perspectives on the Study of Speech*, 1-38.

Cochard, A. (2008). *Jonathan Harvey, Arshia Cont et Gilbert Nouno present Speakings at l'Ircam* [Video file]. Retrieved November 13, 2020, from https://www.youtube.com/watch?v=wjFz-tfCoPk&ab_channel=duduchier

Harvey, J. (1981). "Mortuos Plango, Vivos Voco": A Realization at IRCAM. *Computer Music Journal*, 5(4), 22-24. doi:10.2307/3679502

Harvey, J. (1999). *In quest of spirit: Thoughts on music*. Berkeley, CA: University of California Press.

Fineberg, J. (2000). Guide to the Basic Concepts and Techniques of Spectral Music. *Contemporary Music Review*, 19:2, 81-113. doi: 10.1080/07494460000640271

Harvey, J. (1999) The metaphysics of live electronics. *Contemporary Music Review*, 18:3, 79-82. doi: 10.1080/07494469900640351

Harvey, J. (2008) *Speakings for Orchestra and Electronics*. London: Faber Music.

House, A., Stevens, K., Sandel, T., & Arnold, J. B. (1962). On the learning of speechlike vocabularies. *Journal of Verbal Learning and Verbal Behavior*, 1(2), 133-143. doi:10.1016/s0022-5371(62)80010-3

Jones, D. (1987) Compositional Control of Phonetic/Nonphonetic Perception. *Perspectives of New Music*, 25(1/2), 138-155.

Ladefoged, P. (2006). *A course in phonetics*. Boston, MA: Thomson Wadsworth.

Ladefoged, P. (2010). *Vowels and consonants: An introduction to the sounds of languages*. Malden, MA: Blackwell.

Lisker, L. (1985). The pursuit of invariance in speech signals. *The Journal of the Acoustical Society of America*, 77(3), 1199-1202. doi:10.1121/1.392185

Nouno, G., Cont, A., G. C., & J. H. (2009) Making an Orchestra Speak. *Proceedings of the Sound and Music Computing Conference*, 277-282.

Samuel, A. G., & Tartter, V. C. (1986). Acoustic-Phonetic Issues in Speech Perception. *Annual Review of Anthropology*, 15(1), 247-273. doi:10.1146/annurev.an.15.100186.001335

Schnell, Norbert, and Schwarz, Diemo. (2005) Gabor, Multi-Representational Real-Time Analysis/Synthesis. Proc. of the 8th Int. Conference on Digital Audio Effects.

Smalley, D. (1997). Spectromorphology: Explaining sound-shapes. *Organised Sound*, 2(2), 107-126. doi:10.1017/s1355771897009059

Stevens, K. N., & Blumstein, S. E. (1979). Acoustic invariance in speech production: Evidence from measurements of the spectral characteristics of stop consonants. *The Journal of the Acoustical Society of America*, 66(4), 1001-1017.

Wishart, T. (1996) *On Sonic Art*. Amsterdam: Harwood Academic Publishers.