

ANALYSIS OF ELECTROACOUSTIC WORKS WITH MUSIC AND LANGUAGE INTERSECTIONS

Rodrigo Segnini

Center for Computer Research
in Music and Acoustics (CCRMA)
Stanford University

Bruno Ruviaro

Center for Computer Research
in Music and Acoustics (CCRMA)
Stanford University

ABSTRACT

This paper proposes a theoretical model for the analysis of music and language intersections in contemporary music. This model is tested on a repertoire of electroacoustic and acoustic pieces, providing examples of blurred boundaries between music and speech. We have defined two axes representing possible dimensions of this relationship: intelligibility and music/speech qualities. The resulting sonic space provides a visual account that can serve as a “travel guide” in the music-speech continuum. We envision applications of this model in the fields of composition, musicology, and cognitive studies.

1. INTRODUCTION

The human voice has played a significant role in composition throughout the history of music. Among important issues in the vocal repertoire are text intelligibility and the dichotomy between singing and speaking.¹ Electronic means of music production and delivery have increased the potential for experimentation in this regard, resulting in an increased blurring of the distinction between music and language acting as driving force in many recent compositions.

Normally, music and language are considered separate domains. The following compilation² offers an opportunity to review the categories traditionally implied by the words music and language. Our goal in this work is to formalize, in an intuitive manner, the space in which their intersections occur. We developed a bi-dimensional representation of this space where selected pieces are located, facilitating their comparative analysis. We hope that this may be helpful to musicians and researchers of language and cognition.³

¹ Text intelligibility is not only a linguistic concern, but also a musical one: it contributes to the shaping of different styles and their historical study. For example, one of the major differences between Renaissance and Baroque repertoires is related to the treatment of text. During the Renaissance, there was a tendency to privilege the multi-layering of voices and the use of simultaneous texts in detriment of intelligibility. In opposition to this practice, composers from the Camerata Fiorentina and after prioritized the clarity of the text, which formed the basic principles of the Baroque era.

² Also documented in [5] and [6].

³ [1] and [2] are examples of previous work on the comparison between music and language in the field of music cognition.

2. METHODOLOGY

We have selected 15 pieces that invite fundamental questions about the boundaries between music and speech. Each piece meets the following criteria:

- Vocal sounds in any imaginable form should play a relevant role in it. Synthetic sounds that emulate speech or other vocal behaviors also qualify.
- It should be representative of the various post-tonal musical structuring processes that flourished throughout the 20th century.

| Composer | Piece | * |
|---------------------------------------------------|------------------------------------------------------------|----------|
| SCHOENBERG, Arnold (1874-1951) | <i>Pierrot Lunaire</i> (1912) | 1v + ins |
| SCHWITTERS, Kurt (1887-1948) | <i>Ursonate</i> (1922-32) | 1v |
| SCHAEFFER, Pierre (1910-95); HENRY, Pierre (1927) | <i>Symphonie pour homme seul</i> (1950) | ea |
| STOCKHAUSEN, Karlheinz (1928) | <i>Gesang der Jünglinge</i> (1955-56) | ea |
| BERIO, Luciano (1925-2003) | <i>Visage</i> (1961) | ea |
| EIMERT, Herbert (1897-1972) | <i>Epitaph für Aikichi Kuboyama</i> (1960-62) | ea |
| LIGETI, György (1923) | <i>Aventures and Nouvelles Aventures</i> (1962-66) | 3v + ins |
| LUCIER, Alvin (1931) | <i>I am sitting in a room</i> (1970) | ea |
| DODGE, Charles (1942) | <i>Speech Songs</i> (1972) | ea |
| BERIO, Luciano (1925-2003) | <i>A-Ronne</i> (1974) | 6v |
| APERGHIS, George (1945) | <i>Récitations</i> (1977-78) | 1v |
| LANSKY, Paul (1944) | <i>Six Fantasies on a Poem by Thomas Campion</i> (1978-79) | ea |
| BODIN, Lars-Gunnar (1935) | <i>On Speaking Terms II</i> (1986) | ea |
| MOSS, David (1949) | <i>Direct Sound: Five Voices</i> (1989) | 5v |
| WISHART, Trevor (1946) | <i>Tongues of Fire</i> (1994) | ea |

Table 1: List of musical examples. [*] “v” = voice (live singer or performer); “ins” = acoustic instruments; “ea” = electroacoustic (“tape”) pieces. See [9] to [23].

Six non-electroacoustic pieces were included due to their utility as points of reference for our analysis. In the visual representation that follows we also refer to other genres and two non-Western musical traditions, to help us clarify the scope of our model.

3. ANALYSIS

3.1. Dimensions

In order to gain more insight from this compilation we have defined two dimensions that distill some of its important characteristics in a compact way. We consider these dimensions to be related to modes of perception of an ideal listener. One dimension (3.1.1) represents the degree of text intelligibility in an audio signal, and the other dimension (3.1.2) represents the listener's judgement of the audio signal, according to its speech-like and/or music-like features.

3.1.1. Intelligibility

Intelligibility is contingent on the context of occurrence of a message and apriori knowledge of its construction rules, shared between the producer and receptor. In our case, it relates to the accuracy with which the listener understands the text. This accuracy can be influenced by different processes, deliberate or not, dealing with the message's code, realization and context:⁴

- Code manipulation affects the degree of compliance of the message with linguistic rules (e.g., syntactic or semantic) shared by producer and receptor;
- Realization affects the degree of sonic alteration with which the message is delivered, i.e.: the way the voice is sonically transformed;
- Context relates to other interventions dependent on the circumstances of presentation of the message (different places, external or background noise, etc.).

The first two aspects above can be highly determined by the composer's intentions,⁵ whereas the third is generally less predictable. The interaction of these three factors determines the overall intelligibility perceived by the listener. Taking maximum intelligibility as one of the extremes of a continuum, towards the other extreme the vocal expression becomes more and more linguistically disorganized in ways unexpected by the listener, until it becomes minimally intelligible at the other extreme.

3.1.2. Musicness-Speechness

Another useful dimension extends along a continuum between two other poles of reference: music and speech. It describes the perceived position of a sonic example as either music or speech in a traditional sense, but also comprising all possibilities in between. This, in fact, can be quantified. For example, [7] describes a speech/music discriminator that relies on a number of acoustic features whose quantities correlate strongly with either domain.⁶ The core concept of the dimension *musicness-speechness* is the idea of a continuum between events perceived as music-like or speech-like.

⁴ Here we are referring to "message" as the vocal component of a composition, which may include non-vocal components as well.

⁵ The composer has virtually total control of the first two in the case of tape pieces. In instrumental pieces, performers have different degrees of freedom in their interpretation (especially in the second aspect).

⁶ For example, one of their features is based on the assumption that it is more likely to find an energy modulation peak at around 4 Hz in speech signals than in music signals. Another one relies on the notion that music tends to have a higher spectral centroid than speech. [7]

For example, taken together with the previous dimension, normal conversational speech using meaningful and well-formed grammatical structures carries the highest degree of *speechness* and intelligibility.

3.2. Music-Language Sonic Space

Following an empirical method, we placed some of the selected pieces into the music-language sonic space (Fig. 1) formed by the dimensions above. The placement of a sonic event in a single location represents an average of its most salient features. For events that resist to averaging in a specific context, an arrow is attached to suggest how the different moments seem to navigate through the space. A dynamically changing visualization tool would be useful to show such progressions in time.

A sonic event deserves a unique location in this space if the variance of its characteristics is sufficiently small relative to the other pieces in consideration. For example, if a collection of events using *Sprechstimme* ends up sharing the same location with *Pierrot Lunaire*, a more detailed focus on the internal aspects that make each of them different would be necessary. In this paper, we provide examples that are distinct enough to occupy different sectors of the space, as well as pieces that in average tend to share the same location. In the latter case, they are placed inside a circle to enhance visibility.

3.2.1. Commentary on selected pieces

We assume that the reader is familiar with most pieces listed in Figure 1. We will comment on a few specific examples to facilitate the understanding of the music-language sonic space.

In Alvin Lucier's *I'm sitting in a room*, the speaking voice of a narrator is captured by a microphone and fed back through loudspeakers into the same room. Speech sounds are progressively destroyed by the filtering out of all spectral components that are not part of the natural resonances of the room. After dozens of iterations, speech qualities are effaced, resulting in a sonic texture perceived as music. That is why this piece is seen traversing diagonally the music-language space from the *pure speech/intelligible* to the *pure music/unintelligible* extremes.

Eimert's *Epitaph* and Lansky's *Six Fantasies* are initially placed in the same location as Lucier's, yet their trajectories are not continuous (i.e. different moments of the piece jump from one point to another along the line). In Eimert's piece, for example, sound processing techniques applied to a recorded speech create a sonic palette containing various shades of text intelligibility and music-speech qualities.

A speech signal can be sonically transformed up to the point at which even the presence of a human voice becomes unrecognizable—this point represents the extreme of the *musicness* side, a definitely non-vocal timbre region. The fact that some pieces can share the same location and path, despite their differences, shows that there exist many musically and technically different ways of traveling the music-language space.

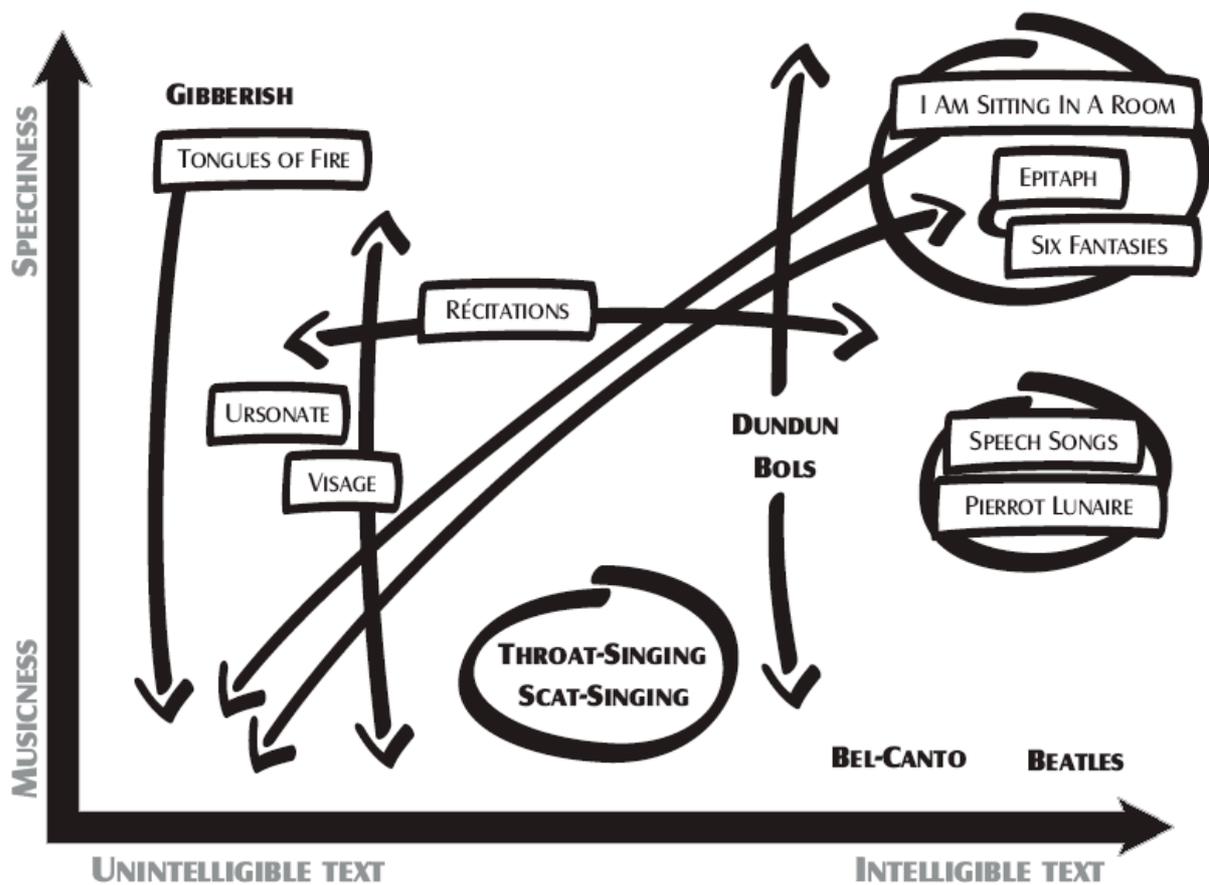


Figure 1. Some examples from Table 1 placed within the music-language sonic space.

We have placed in Figure 1 some additional sonic events (not enclosed in boxes) to be used as references: *gibberish* is located at the top-left corner for its unintelligibility, yet usual characterization as speech. *Scat-singing* and *throat-singing* appear halfway on the intelligibility axis due to their usual lack or dissolution of verbal meaning, and towards the music extreme for their status as musical manifestations. *The Beatles* are placed on the lower-right corner, as a prototype of music with a very intelligibly sung text, while *bel-canto* is located to its left.

Finally, consider two examples borrowed from non-Western cultures: the *Yoruba* drums and the *Tabla bols*. *Yoruba* is an African tone language with 3 intonations: low, medium and high. Words are basically differentiated by particular sequences of tones, which the *Yoruba dundun*⁷ musical practice is able to convey. The *bol system* is a collection of vocal utterances that mime the sounds produced by the *tabla*.⁸ In both cases, the spoken and musical domains are intertwined, even for people unfamiliar with these cultures.⁹

⁷ Set of double-headed hourglass tension drums. [1]

⁸ The *tabla* is an asymmetrical pair of small, tuned, hand-played drums of North and Central India, Pakistan and Bangladesh. [1]

⁹ The universality of this kind of onomatopoeic effect was studied in [4] for the *Tabla* case.

3.3. Limitations and Future Work

This work presupposes a listener who is willing to look for the *musicness* in speech, or the *speechness* in music. Such an active listener is able to put aside traditional prejudices that confine a sonic event into either one of the categories of speech and music.

Obviously, this ideal listener knows the codes being used; unfamiliarity with the language would result in judgements of unintelligibility as well as a heightened musical perception of a message otherwise recognized simply as speech.

Experimental studies are being conducted to test the validity and the implications of the chosen dimensions.

4. CONCLUSIONS

Any piece can theoretically traverse the entire space according to the composer's imagination. However, because of the intrinsic limitations of the human voice as well as of the acoustic instruments, this potential could not be fully exploited until the advent of electroacoustic music technology. The means of sound production, transformation and delivery developed in the last fifty years expanded the ability of composers to go beyond those limitations. For example, Eimert's *Epitaph* is one of the first conscious attempts to explore

the ambiguities of the speech-music continuum using some of the earliest electronic means.

We believe that the presented space can be used by composers interested in exploring music-speech intersections in a more formalized way.¹⁰ There are, according to this particular selection of pieces, uncharted territories and spatial movements that suggest opportunities for aesthetical exploration. Naturally, those locations already “colonized” by other pieces may also be exploited through different means, resulting in different pieces.

Another application is to use this space as a framework for the development of new experiments in music-language cognition. In general, most musical material used in current research belongs to the Western common practice or popular music. The multiple possibilities of blurring the boundaries between music and language, as demonstrated in the pieces selected, may provide new and useful insights for our understanding of the nature of perception in both domains.

5. REFERENCES

- [1] Aiello, R. “Music and Language: Parallels and Contrasts”, in Aiello, R. & Sloboda, J. (Eds.), *Musical Perceptions* (pp. 40-63). Oxford University Press, New York, 1994.
- [2] Besson, N. & Schön, D. “Comparison between Language and Music”, in Zatorre, R. & Peretz, I. *The Biological Foundations of Music* (pp. 232-256). The New York Academy of Sciences, New York, 2001.
- [3] *Grove Music Online* ed. L. Macy. Accessed [March 7 2005], <http://www.grovemusic.com>).
- [4] Patel, A. and Iversen, J. “Acoustic and Perceptual Comparison of Speech and Drum Sounds in the North Indian Tabla Tradition: An Empirical Study of Sound Symbolism, *Proceedings of the 15th ICPHS*, Barcelona, Spain, 2003.
- [5] Ruviaro, B. *The Spell of Speech*. Master’s Degree Thesis defended at Dartmouth College (New Hampshire). Hanover, USA, 2004.
- [6] Ruviaro, B. and Segnini, R. “Music and Language Intersections in 20th Century Post-Tonal Music: Overlooked Opportunities for Research in Cognitive Neuroscience” (submitted to the *II Conference on Neurosciences and Music*, Leipzig, 2005).
- [7] Scheirer, E. and Slaney, M. “Construction and Evaluation of a Robust Multifeature Speech/Music Discriminator”, *Proceedings ICASSP-97*, Munich, Germany, 1997.
- [8] Segnini, R. “Music Score Phonetization and Speech-based Composition”, *Proceedings of the International Symposium on Musical Acoustics*, Nara, Japan, 2004.

5.1. Recordings

- [9] Aperghis, G. *Récitations*. Sne Records, 1997, CD 571.
- [10] Berio, L. *Visage*. Ricordi, 1991, CD 1017.
- [11] Berio, L. *A-Ronne*. RCA, 1998, CD 68302.
- [12] Bodin, L. *On speaking terms II*. Available at www.ubu.com
- [13] Dodge, C. *Speech Songs*. New Albion Records (San Francisco), 1992, CD NA 043.
- [14] Eimert, H. Epitaph für Aikichi Kuboyama. Wergo (Germany), 1962, WER 60014.
- [15] Ligeti, G. *Aventures and Nouvelles Aventures*. Sony Classical (New York), 1997, SK 62311.
- [16] Lucier, A. *I am sitting in a room*. Lovely Music Ltd., 1990, CD 1013.
- [17] Lansky, P. *Six Fantasies on a Poem by Thomas Campion*. CRI, 1994, CD 683.
- [18] Moss, D. *Direct sound: Five Voices*. Intakt, CD 015.
- [19] Schaeffer, P. & Henry, P. *Symphonie pour un homme seul*. Schaeffer Complete Works, INA/GRM (France), CD 1006/7/8.
- [20] Schoenberg, A. *Pierrot Lunaire*. Ensemble InterContemporain; Pierre Boulez, conductor. Sony, 1993, CD 48466.
- [21] Schwitters, K. *Ursonate*. Available at www.ubu.com
- [22] Stockhausen, K. *Gesang der Jünglinge*. Stockhausen Verlag (Germany), 1991, CD 3.
- [23] Wishart, T. *Tongues of fire*. EMF, 2000, CD 029.

¹⁰ As in [5] and [7].