The Maximal–Minimal Model: A framework for evaluating and comparing experience of voice in electroacoustic music

ANDREAS BERGSLAND

Department of Music, Norwegian University of Science and Technology (NTNU), 7491 Trondheim, NORWAY
E-mail: andreas.bergsland@ntnu.no

This article presents a framework for describing, understanding and evaluating the experience of voice in acousmatic electroacoustic music and related genres through the *maximal–minimal model*. This model, which is inspired by literary theory, theories of radiophonic voice as well as theories of electroacoustic music, presents *maximal* and *minimal* voice as loosely defined poles constituting end points on a continuum on which experienced voices can be localised. Here, *maximal* voice, which parallels the informative and clearly articulated speaking voice dominant in the radio medium, is described as the converging fulfilment of seven *premises*. These premises are seen as partly interconnected conditions related to particular aspects or features of the experience of voice. At the other end of the continuum, *minimal* voice is defined as a boundary zone between voice and non-voice, a zone which is related to the negative fulfilment of the seven premises. The two poles are presented as centre and periphery, respectively, with the seven premises constituting multiple axes spreading out from the centre. These features, it is argued, parallel Lakoff's cluster model of categorisation. Lastly, the article briefly discusses the use of the framework in analysis of electroacoustic works with voice, and it demonstrates two ways in which the evaluations according to the framework can be visualised.

1. INTRODUCTION

When listening to, interpreting and analysing electroacoustic pieces with voice, aspects of meaning and provenance are often of central importance. While several scholars have written with great insight on this topic (see for example Wishart 1996: part 3; Norman 2000; Bossis 2005), there still seems to be a lack of a framework for describing, understanding and evaluating listeners' experience of voice in pieces of electroacoustic music in a systematic manner. The framework presented in my PhD thesis *Experiencing Voices in Electroacoustic Music* (Bergsland 2010) was an attempt to remedy this.[1] As the title suggests, this framework has a phenomenological basis, where an investigation of my own listening process through internal subjective inspection has been the most important methodological tool. Moreover, my aim of seeing how the experience is formed by previous experiences, knowledge and predilections and the way in which I orient my attention and consciousness during listening is similar to other approaches that are explicitly phenomenological (e.g. Ferrara 1984).

The first major part of the framework comprise a set of *experiential domains* pointing to a class or group of properties assigned to an object of the experience and based on some common feature, relationship or function. However, the focus in this article will be put on the second part of the framework put forward in my thesis, namely the *maximal–minimal model*. This model sets up two poles or extremes as reference points against which the experience of more or less transformed or manipulated voices might be judged and compared; *maximal* and *minimal* voice. The maximal voice can briefly and preliminarily be described as a typical informative and neutral speaking voice, resembling in many ways a public radio broadcast voice. At the other end, the minimal voice is usually highly manipulated and often quite abstract, and thus defines the zone between what *is* voice and what *is not* voice. The imagined space between these two extremes is thought of as a continuum extending from a central zone, defined by the maximal voice, towards a peripheral zone, defined by the minimal voice. This continuum is also mapped out and described in a more detailed manner by formulating a set of *premises*, which can be thought of as partly interrelated properties or dimensions with which different vocal expressions in electroacoustic music can be evaluated and/or compared. Both the centre-periphery and the multi-premise ideas will be discussed in more detail in the sections below.

Even if the *maximal–minimal model* is based on the *experiential domains* part of the framework, I will attempt to describe the former without reference to the latter, and the interested reader should consult my dissertation on the details of the framework as a whole (Bergsland 2010).

[1]An earlier version of the framework can also be found in Bergsland 2006.

Organised Sound **18**(2): 218–228 © Cambridge University Press, 2013. doi:10.1017/S1355771813000137

2. BACKGROUND

The concepts of maximal and minimal voice are borrowed from the literary theorists Donald Wesling and Tadeusz Slawek and their use of these concepts in the book *Literary Voice: The Calling of Jonah* (Wesling and Slawek 1995). The two concepts are presented as ends or extremes of literary voice, and these are theorised by the authors by analysing texts that are seen as belonging to each of the extreme positions. Although literary voice is a more abstract notion compared to audible voice, fundamentally bound up in the often highly problematic relationship between speech and writing, there are several issues addressed by these authors that appear to be relevant also for loudspeaker mediated voices, like meaningfulness, intelligibility, presence, the relationship between voice and person, expressivity, intentionality, individuality and bounding outline. Many of these issues can be recognised in the seven premises of the *max–min model*, even if they have had to be greatly adapted and re-worked.

In addition to Wesling and Slawek, two radio theorists have presented ideas that have had an importance for the development of the model. Firstly, Bodil Børset, also drawing upon Wesling and Slawek, has analysed Nathalie Sarraute's *pièces radiophoniques*, focusing on aspects such as the strength of the link between character and voice, the degree of separation between a voice and background noise, voice familiarity, intelligibility, the 'materiality of sound', and the dissolution of sense and meaning (Børset 2006: 128–31). Here the first four aspects were discussed as factors pulling the voice towards the maximal, whereas the latter two were seen as drawing it towards the minimal. Secondly, what Frances Dyson has written about radio voice corresponds in many ways to the ideas of Wesling and Slawek. In her article 'The Genealogy of the Radio Voice', she mainly focuses on what she calls the 'dominant' radio voice (Dyson 1994). This type of voice has many parallels with the literary maximal voice, even if it doesn't explicitly relate it to the idea of a continuum and the existence of an opposite pole. She writes that '[generally], the dominant radio voice talks – its speech is clear, articulate, sometimes eloquent. Most of what it says is perceived by the listener as factual and informative … It does not mumble or stutter, it pronounces full and meaningful sentences, it says something' (1994: 167). *Presence* is also an important aspect of Dyson's notion of the dominant radio voice: 'For the voice being sonorous, necessarily speaks in the present, is accompanied by the actual bodily presence of the speaker, and is heard (perceived) at the moment of its production' (1994: 176). Furthermore, she states that a voice also has to be *singular* – in other words, it must appear one at a time, and it has to appear as the only audible sound source, or alternatively as the foreground phenomenon with very few other sound sources interfering, to aspire to the dominant radio voice (1994: 181–2). Despite the fact that Dyson mainly defines the dominant radio voice positively, she also demonstrates ways in which some aspects can be negated, and thereby it is possible to relate these negations to the minimal voice. For instance, she sees both bodily 'noise' – for example trembling voices, throat clearing, coughing, sneezing, panting – as well as noise from the technology of mediation as factors that undermine the dominant radio voice. Taken together, Børset's and Dyson's writings present several issues that are highly relevant for voices in electroacoustic music, and therefore many of them are adapted to and incorporated in the framework; that is, in the seven premises that I will present below.

The max–min model also draws on theories of electroacoustic music, both more general theories and ones specific to voice. As for the former, I have been influenced by Smalley's theory of transformation, and in particular his idea of source-bonded transformations as relying on an *extrinsic identity* – in other words, a mental image of a general or specific sound source or action that exists outside of the work (Smalley 1993: 280–1). More specifically, the way he conceives of *transformational distance* as a 'subjective measure of degree of change of identity relative to the base identity' (1993: 287) has been particularly influential. This parallels how I regard the experience of the (often transformed) voices in electroacoustic music as subjectively defined through the relationship to the (untransformed) *maximal voice*, and how this interrelationship can be expressed and visualised spatially – as located on a continuum closer or further from the *maximal voice*. I will return to the details of this below.

Furthermore, the maximal–minimal model also draws upon Emmerson's notion of *physical presence*, which deals with listeners' reconstruction of physical action and agencies, including all types of (living) sound sources and actions, and their disposition in and interaction with an environment (Emmerson 2007: 18–23). Finally, I am also relying on theoretical models that have described the continuum between the concrete and reference oriented on one side, and the abstract and sound-quality oriented on the other side (Chion 1988; Emmerson 1986; ten Hoopen 1992; Young 1996).[2]

The theories that are specific for voice in electroacoustic music are naturally even more interesting.

[2]I see this idea as a question of what aspects of the listener attention is focused on, and, as I argue in my thesis, it can therefore be regarded as incorporated in the *focus of attention* premise (see Bergsland 2010: chapter 5).

https://doi.org/10.1017/S1355771813000137 Published online by Cambridge University Press

In Bruno Bossis's groundbreaking book *La voix et la machine* he introduces the concept of *artificial vocality*, a term that in its turn conjoins the concepts of 'vocality' and 'artificiality', in principle covering all aspects of vocal or vocal-like expressions in sound mediated by loudspeakers, but in practice focusing on voice in electroacoustic works (Bossis 2005). By using the term 'vocality' rather than speaking of the voice, Bossis opens up for seeing the phenomenon as a graded quality rather than a question of either–or. In summing up the ideas presented in the book, Bossis emphasises how artificial vocality is played out in a generalised continuum, where all musical parameters are continuous (2005: 288). More specifically, he draws up a continuum between the 'frankly vocal' ('franchement vocal') and what is 'not at all vocal' ('pas du tout vocal') (2005: 288). What is more, he argues that the detachment of the voice from bodily production, which characterises artificial vocality, delineates a continuum between the human and the synthetic (2005: 289). All in all, the idea of the continuum is a very important one in Bossis's discussion of artificial vocality, and it is not difficult to see the similarity between these ideas and the gradation between the maximal and the minimal fulfilment of the *naturalness* premise.

A much less developed theory, but still interesting in this context, is presented by Segnini and Ruviaro in their paper 'Analysis of Electroacoustic Works with Music and Language Intersections' (Segnini and Ruviaro 2005). Here, the authors break down the analysis into two dimensions in what they call a 'music-language sonic space', where 'musicness <–> speechness' constitutes one axis and 'unintelligible text <–> intelligible text' the other. These dimensions resemble certain of the premises of the model I will present below, and their use of a spatial model with two continuous dimensions parallels the way the premises can be treated as dimensions, as will be clear from the forthcoming discussion.[3]

There are also two other theoretic fields that in themselves are perhaps not related to Wesling and Slawek's ideas of literary voice, but that I still have found relevant for my framework, namely cognitive load theory (Halford, Wilson and Phillips 1998; Paas, Renkl and Sweller 2004) and other related theories of information and redundancy (Eco 1989), as well as auditory scene analysis theory (Bregman 1990). Again, it would be outside the scope of this article to deal with these theoretic fields here, and readers should consult chapters 6 and 12 of my thesis (Bergsland 2010) for details.

[3]I am thinking here of *dimensions* not in a strictly mathematical sense, but as relative and inexact axes that correspond to the gradation along a continuum of one fairly coherent aspect of an experience.

3. MAXIMAL VOICE AS A SET OF PREMISES

Retaining many of the ideas presented above, I will now formulate the maximal voice as a set of *premises*. The premises each express one particular aspect or feature of the maximal voice, similar to several of the aspects mentioned in the discussion above. Each of these premises can be seen as conditions that can be fulfilled to different degrees, and when they are all fulfilled, the result is what I define as maximal voice. Since fulfilment is graded, the premises can be seen as continua running from the maximal to the minimal, in line with the idea of a graded continuum that we could find both with Wesling and Slawek and with many of the electroacoustic theoreticians. And they can be taken as the basis for an evaluation that can be represented graphically along a set of *axes*, as I will show in section 6 below.

It is possible to conceive of several ways in which the main ideas presented above could be organised into premises, for instance regarding their number, what ideas they comprise and what they are called. And, indeed, I have operated with several set of premises during the work with the model (see Bergsland 2006). Hence, the seven premises that I present here are not 'written in stone' but constitute what I have found to be the clearest and most useful organisation that allows for sufficient distinction and a good balance between overview and detail. Here, I present them together with a set of associated keywords that links them to the theoreticians and theoretical groundings that I discussed in the previous section:

1. **Focus of attention**: The semantic issues of any verbal structures receive sustained and maximal attention. (See sound materiality (Børset 2006); reality–abstraction continuum (Young 1996); musicness–speechness (Segnini and Ruviaro 2005).)
2. **Information density**: Information density is optimal for the processing or decoding of the verbal structures. (See cognitive load theory (Halford et al. 1998; Paas et al. 2004); information, redundancy (Eco 1989).)
3. **Naturalness**: The sound has maximal resemblance with one produced by a human being and his or her vocal apparatus. (See artificial vocality (Bossis 2005).)
4. **Presence**: The listener experiences a sense of a shared 'here and now' with a vocal persona. (See presence (Dyson 1994; Wesling and Slawek 1995; Emmerson 2007).)
5. **Clarity in meaning formation**: Meaning can be constructed from the voice with a high degree of clarity – also implying specificity, certainty and coherence. (See intelligibility, meaningfulness (Wesling and Slawek 1995); intelligibility, voice familiarity, dissolution of sense and meaning (Børset 2006); clear speech, articulate, factual,

https://doi.org/10.1017/S1355771813000137 Published online by Cambridge University Press

informative, full and meaningful sentences (Dyson 1994); intelligible text (Segnini and Ruviaro 2005).)

6. **Feature salience**: Vocal sounds and features 'stand out' perceptually – for themselves and relative to other sounds and features. (See individuality and bounding outline (Wesling and Slawek 1995); separation between a voice and background noise (Børset 2006); singularity, foreground phenomenon (Dyson 1994).)
7. **Stream integration**: The sound of the voice is integrated into one coherent and continuous sound stream. (See auditory scene analysis (Bregman 1990).)

In my thesis, these seven premises are individually given a theoretical grounding and related to a set of *factors*, which affect the evaluation of each of them. This provides the basis for detailed criteria for evaluation of each premise along the max–min continuum, where the importance of the listeners' background, knowledge and experience as well as the listening conditions (e.g. how many times a piece is listened to, or if it is listened to in one stretch or paused) is emphasised. These criteria are then applied in an evaluation of selected excerpts of music. Moreover, I show in my thesis how several of the premises are partly interconnected in that a premise can constitute a factor affecting the evaluation of another premise.

While the details of each premise and its use in evaluation have to be left out in this context, I will instead focus on more general issues. If we start with maximal voice, it is not encountered in its fullest sense too often in the type of material that I have focused on in my work, but tends to be the rule in other sonic expressions as radio and audio-book genres such as the interview, speeches and lectures. One example from electroacoustic music that approaches maximal voice can be found in *Les objets obscures* (1991, on Parmerud 1994) by Åke Parmerud. In this piece, a French female speaking voice is heard at several salient points in the piece. Several times this voice appears on its own, close and without any reverberation, devoid of ambient noise, speaking in an articulate and fluent manner, clearly intelligible, at least for listeners who understand French. Yet on the semantic level there is not so much clarity: the sentences are not always complete, and what is referred to remains mostly rather obscure since the woman is presenting a riddle, a riddle which hints at the 'hidden objects' that can be heard in the piece, and that the title refers to:

> Le deuxième: Un paysage ambulant. Un déplacement perpétuel. Quelque chose qui frôle sans toucher. Un mouvement sans but. Un objet de repos.
>
> (The second: a landscape on legs. A constant moving. Something that touches without touching. A movement without goal. An object to rest in.) (Liner notes from Parmerud 1994)

Even if the answers to the riddles are presented in the fourth part of the piece, the general impression is that this is quite far from the meaningfulness of the informative and factual radio voice. Moreover, the voice is only present in short sections at a time, sections which function mostly as introductions to the more 'musical' parts of the piece. This woman is therefore felt to retreat to an undefined absence during most of the piece, thus being only temporarily present for the listener. One can see that even if this voice doesn't bear all of the characteristics described above, it still seem to largely fulfil many of the premises such as *focus of attention* (at least for those understanding French), *information density*, *naturalness*, *presence*, *salience* and *stream integration*.

4. CENTRE AND PERIPHERY

When the experienced voices only partly fulfil the premises of the model, they will depart from the maximal, and at some point of negative fulfilment, in what we might call a *violation* of the premises, they will ultimately reach a state where they can be characterised as *minimal*.

The minimal voice is much more difficult to define than the maximal voice because it comprises a wider range of possibilities and modes of expression. For the mentioned seven premises, there are simply more ways in which they can be violated than be fulfilled. Thus, it seems that the relationship between the maximal and the minimal can be described through the dichotomy narrow–broad. For the *clarity of meaning formation* premise, for instance, the maximal voice will be confined by conventions of speech and language. The minimal voice, on the other hand, isn't confined by any mode of voice at all – it can comprise speech, singing and vocal experimentation. For example, in Western classical *bel canto* song, which often has a basis in a meaningful text, the text is often unintelligible due to melismatic passages and phrases in high registers that make pronunciation difficult.[4] In artistic vocal expressions of the avant-garde, within as well as outside of music, there are many examples of other kinds of voices that in different ways have presented voice with little clarity of meaning, for example by using non-sense phonetic texts or pseudo-language (as in *Ursonata* by Schwitters (1922–32, on Schwitters 1992) or *Nouvelles aventures* by Ligeti (1962–65, on Ligeti 2006)). And, clearly, in electroacoustic works one can find a great many examples of types of electronic processing that might affect the clarity of meaning negatively: for example,

[4]The max–min model does not account for the way in which conventionalised modes of singing constitute other 'centres' that other types of vocal sounds may be experienced in relation to. As I see it, it would complicate matters too much to include this into the model.

https://doi.org/10.1017/S1355771813000137 Published online by Cambridge University Press

filtering, time stretching and compressing, granulation and distortion.

In addition to representing one of two extremes in my framework, the minimal voice also represents the possibility of transgressions into what *is not voice* – it represents a boundary zone where the voice appears to be on the verge of turning into what is no longer a vocal sound. And, in electroacoustic music the boundary between what is voice and what is not voice is the subject of exploration in several works, among them Stockhausen's *Gesang der Jünglinge* (1956, on Stockhausen 2001), Wishart's *Red Bird* (1977, on Wishart 1992), *Mortuos Plango, Vivos Voco* by Jonathan Harvey (1980, on Various Artists 1990) and *Chant d'ailleurs* by Alejandro Viñao (1992, on Viñao 1994). In these and several other pieces one can experience gradual transformations between vocal sounds and sounds that are clearly of a different origin.[5] In the very beginning of the latter work (0′00″–0′30″), for instance, there is a sustained note which transforms gradually and continuously from a wind instrument with a rather ethnic and Eastern flavour into a singing voice. This happens rather slowly and in several stages, and at one point (I experience this around 0′20″) the sound begins to take on the qualities of a voice. At this point, however, it is still lacking greatly in naturalness compared to a real voice due to the fact that it lacks the small irregular fluctuations in pitch and amplitude that are characteristic of the human voice.[6] When the pitch fluctuations set in a little later, however, the sound becomes more natural, until it finally sounds just like a sung note. One can thereby experience that the sound gradually changes from non-voice into voice, and that, at one point, the sound passes through a boundary zone between the two. I will return to a more in-depth discussion of such boundary transgressions below.

One useful way of visualising the discussed relationships between the maximal and the minimal voice is a circular centre–periphery model (see Figure 1). Here, the maximal voice constitutes the centre and the minimal voice the periphery, which borders onto what is not voice.[7] Moreover, in this model the maximal voice is clearly more narrowly defined than

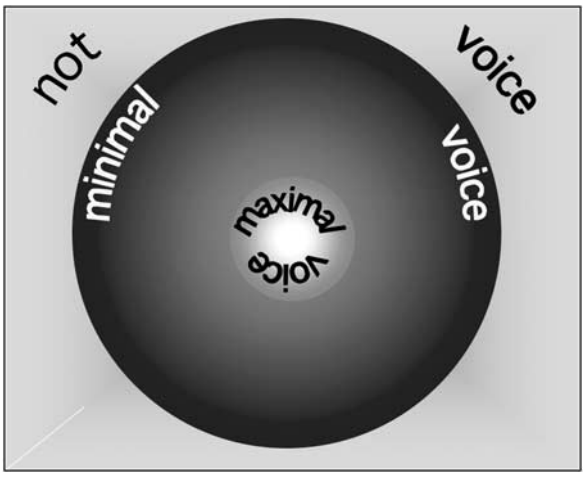

Figure 1. Centre–periphery model of maximal and minimal voice.

the minimal, representing the multitude of ways that voice can depart into the minimal and ultimately into non-voice. The idea of a graded continuum between the maximal and the minimal is also retained in this model. Still, the model lacks the connection to the seven premises introduced above, and I will therefore look into ways of expanding it. Theories on categorisation and prototypes appear to offer such a link, and consequently I will look into these theories in the following section.

5. PARALLELS WITH CATEGORISATION MODELS

Eleanor Rosch's and George Lakoff's writings on categorisation and prototype theory present ideas that pose some interesting links to the framework I have hitherto delineated. I will argue in the following that my model in many ways is structured like a particular type of category described by Lakoff, where the category in my specific case will be 'experienced voice in electroacoustic music'.

Prototype theory was formulated by Eleanor Rosch in the 1970s, but has several predecessors, within philosophy, cognitive anthropology and social psychology (Rosch 1975; Rosch and Mervis 1975; Rosch 1978).[8] Her theory was formulated in opposition to the 'classical' view of categories, which saw them as rooted in the objective structures in the physical world. Within the classical view of categories, certain members of a category could *not* be seen as more central or typical to the category than others, since the categories in themselves would be

[5]See, for example, Smalley 1993, Landy 1993 and Wishart 1996 for accounts of transformations or metamorphoses in electroacoustic music.

[6]It has been shown that random variations in pitch and amplitude, so called *jitter* and *shimmer*, are factors that have a positive effect on the evaluation of naturalness of synthetic voices. See, for example, Aoki and Ifukube 1996 or Verfaille, Gustavino and Depalle 2005.

[7]The idea of the maximal voice as a centre is not explicitly formulated by Wesling and Slawek, but can in my opinion be a consequence of their idea of a maximum pole or end from which departures towards the margins and the two minimal modes take off in different directions.

[8]Lakoff mentions Ludwig Wittgenstein's concept of *family resemblances*, Brent Berlin and Paul Kay's work on colour terms, and Roger Brown's study on basic-level categories as important predecessors for Rosch's prototype theory. See chapter 2 of Lakoff 1987.

https://doi.org/10.1017/S1355771813000137 Published online by Cambridge University Press

determined by shared properties rooted in an external and objectively given world, thus endowing all members of a category equal status. However, through reviewing a series of earlier empirical studies and undertaking a set of new ones, Rosch was able to find what she called 'prototype effects'; in other words, that some members of a category were taken to be more prototypical or better examples of the category than others. For example, she found that some species of birds were thought of as better examples of the category 'bird' than others; whereas robins and sparrows were considered the best examples, owls and eagles were not so good examples, and penguins, emus and ostriches were considered worst examples of the category (Rosch 1975).[9]

If we go a bit further into the claims of prototype theory and related theories, we can see that the relationship between the typical and the less typical members also resembles the relationship between maximal and minimal voice, if we regard the whole continuum between the poles as constituting the category 'experienced voice in electroacoustic music'. In the same way as minimal voice is defined in relation to maximal voice, in prototype theory non-prototypical members of a category are defined *in relation* to the prototypical. In the words of Lakoff and Johnson, '[we] understand the nonprototypical chairs as being chairs, not just on their own terms, but by virtue *of their relation to a prototypical chair*' (Lakoff and Johnson 1980: 122, my italics). If we return to the premises that define the maximal voice, we can see that they at the same time define what the minimal voice *is not*. Thus, anything that would be considered as minimal voice would be, in a similar manner as in prototype theory, defined in relation to the maximal voice. In this respect, my model seems to fit the structure of categories in prototype theory.

Moreover, one can see that the spatial metaphor that is implied for the prototypical categories, with some members that are more central than others, in many ways resemble my model, in that they both can be thought of as graded continua between centre (the prototypical members) and periphery (the worst examples). This is evident from Lakoff's summary of the basic results of prototype theory: 'Some categories, like tall, man or red, are graded, they have inherent degrees of membership, fuzzy boundaries, and central members whose degree of membership (on a scale from zero to one) is one … Other categories, like bird, have clear boundaries; but within those boundaries there are graded prototype effects – some category members are better examples of the category than others' (Lakoff 1987: 56). Without addressing the question of the different kinds of boundaries mentioned here at this point, we see that in both cases there is a question of *graded prototype effects*.

As we saw above, I listed several constituent premises that together converged to define the maximal voice of my model. This resembles what Lakoff calls *cluster models*, which designate a source of prototype effects in his theory. Cluster models involve several *cognitive models*, which for Lakoff are mental constructs involved in the organisation and structuring of knowledge and meaning: for example, in forming categories (Lakoff 1987: 74). These cognitive models will in some cases cluster together or converge to form categories that are psychologically more basic than the models taken individually, hence the term cluster models. When all the cognitive models in a cluster converge, it will then result in a more central or prototypical category member than when there are only just a few models clustering or no clustering at all, something which will result in more peripheral members.

Hence, we can see that this way of thinking about categories appears to have many similarities with the way in which the maximal–minimal framework was structured; the maximal voice was defined according to a set of premises, in many ways resembling the cognitive models of Lakoff; the minimal voice was defined *in relationship* to the maximal; and the maximal was seen as more central than the minimal.[10] One difference, however, is that, whereas the subcategories in Lakoff's case were restricted to established, conventionalised categories, my model is in principle open to all kinds of combinations of fulfilment or violation of the premises.

This resemblance with Lakoff's structuring into *cluster models* seems to reinforce the spatial representation in a circular centre-periphery earlier shown in Figure 1. What is more interesting is that it seems to open up for an inclusion of the premises of the model into this picture, since the premises are seen as converging in the centre and diverging towards the periphery. When also taking into account that the premises can be regarded as graded continua running from the centre to the periphery, the spatial representation of the premises as kinds of dimensions having axes pointing outwards from the centre, almost suggests itself. Consequently, an expanded

[9]Another example is referred to by Lakoff and Johnson, whose theories have been greatly influenced by prototype theory: 'A prototypical chair, for us, has a well-defined back, seat, four legs, and (optionally) two armrests. But there are nonprototypical chairs as well: beanbag chairs, hanging chairs, swivel chairs, contour chairs, barber chairs, etc.' (Lakoff and Johnson 1980: 122).

[10]Regarding the prototypicality of the maximal voice, however, it can be questioned in what respect and on what grounds maximal voice can be considered as prototypical or as a 'best example' of the category 'experienced voice in electroacoustic music' apart from sharing the graded centre–periphery structure. See Bergsland 2010: 152–3 for a further discussion.

https://doi.org/10.1017/S1355771813000137 Published online by Cambridge University Press

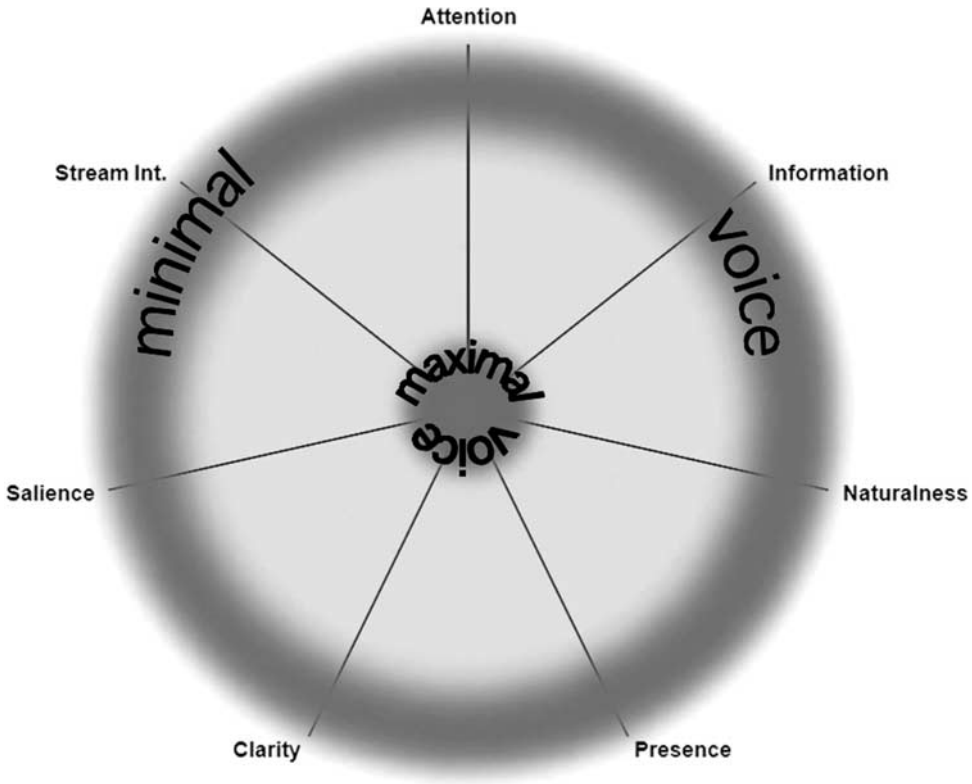

Figure 2. Extended centre–periphery representation of the maximal–minimal framework. The seven axes represent the seven premises introduced in section 3 of the article.

version of the centre-periphery model would be something like Figure 2.[11]

However, I want to emphasise that despite the similarity with mathematical representations of an n-dimensional space, the axes in this case are not 'true' dimensions in the sense of being orthogonal or independent of each other.[12] Rather, I think of this representation as a system of axes in the sense proposed by Godøy (1997). Godøy regards the notion of axes as applicable in principle to anything that can be regarded as an aspect of music, and axes are seen as having both a hermeneutical role as a visualisation and understanding of an aspect, and a role in generation or simulation (1997: 186–7). While the latter is not an issue in this context, Godøy's view of the generality of axial representation in representing aspects along a continuum seems to fit perfectly with my model, and his emphasis on the hermeneutical role that allows for understanding and visualisation is also very much in line with my intention of using such a representation.[13]

While formalisations as the one proposed here always risk simplifying matters and falling short of representing the singularities of individual pieces and listenings, there are also important benefits. Perhaps the most important point of these axes is that they allow for *comparisons*, both between two or more different segments of music and between single segments and hypothetical values along an axis: it is by seeing musical segments, both actual and hypothetical ones *in relation to one another*, that the possibility of knowledge and understanding lies, according to Godøy (1997: 190). Thus, by assigning a segment of music to a value along one axis, one has the possibility of seeing this segment *in relation to* the other values on the continuum, for instance the maximum or the minimum, as well as the possibility of seeing it in relation to another segment of music evaluated with a value along the same axis. Hence, axes in Godøy's sense are tools for relational thought, rather than an absolute mapping of exact values where the assignment of a value along an axis for a musical segment constitutes an act of relative comparison.

[11]One important difference between Figure 1 and Figure 2 that must be noted is that I have removed the 'non-voice' label. This is done because not all of the premises can transcend from the minimal voice to non-voice, consequently making the questions of boundary transgressions more complex and more difficult than for the simple centre–periphery model. See Bergsland 2010: section 4.7 for a discussion of the questions of boundaries, also relating that question to theories of categorisation.

[12]A true dimensional model in the mathematical or physical sense would require full independence between the dimensions or axes of the model.

[13]A similar way of using multiple axes has also been used by Birnbaum, Fiebrink, Malloch and Wanderley 2005.

https://doi.org/10.1017/S1355771813000137 Published online by Cambridge University Press

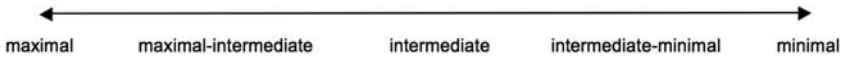

Figure 3. Five value categories along the maximal–minimal continuum.

This value might be graded in different resolutions, from coarse (high–medium–low) to fine (1–20), depending on the possibilities for differentiation along an axis (1997: 146). I largely apply a relatively coarse resolution with the continuum divided into five categories, of which two are the maximal and the minimal poles, one is at the intermediate position, and the remaining two are between the intermediate and the maximal and minimal poles (see Figure 3). In principle, however, one can choose other resolutions depending on the degree to which it is possible to make distinctions and comparisons.

6. ANALYTICAL APPLICATIONS

The maximal–minimal model, supported by the experiential domains framework mentioned in the introduction, can also be applied analytically dealing with concrete works of music. The criteria mentioned in section 3 provide a way of evaluating phrases from particular pieces according to all the seven premises (Bergsland 2010: chapters 5–11), and the result of such an evaluation can then be visualised in two ways – namely, *axial* and *time-varying* representations – which I will demonstrate in the following two subsections.[14] In this article, I will explain these types of representations using the first 27 seconds of the third movement of Paul Lansky's *Six Fantasies on a Poem by Thomas Campion* entitled *her reflection* as an example (Lansky 1979, on Lansky 1994).[15]

6.1. Axial representations

In Figure 4, I have shown the two types of representation along with a traditional spectrogram and a segmentation of the different phrases, the latter shown as rounded rectangles with text grouped into four different phrase categories (a, b, c, d) based on a similarity/difference evaluation, all made with the use of the Acousmographe software.[16] If we start to take a look at the axial representations, they are located in the chain of white rectangles at the top of the graphical panels. We can see here that the shapes are based on the extended centre–periphery representation with the seven axes presented in Figure 2 loosely combined with the five value categories presented in Figure 3. In all representations, the fuzzy circular grey area in the centre represents the zone of the maximal voice, and the peripheral fuzzy circular grey zone represents the zone of the minimal voice. These areas are drawn rather fuzzy or diffuse to underline the lack of any precisely defined limit for these evaluations. The evaluation of each of the premises at one particular time, indicated by a vertical grey line, is then represented by points on each of the seven axes. By defining a line between each of the points and an area enclosed within this line, one will have a representation which shows the evaluation for all the different premises together, especially since the *shape* and the *size* of the area give a lot of information just at one glance. For example, in cases where the area approaches a circular shape with the crossing of the axes at the centre, one will have evaluations which are relatively similar for all of the premises. Conversely, if the area constitutes a shape with many edges and irregularities, this indicates that the evaluations are more dissimilar, thus covering a greater span between the minimal and the maximal. As for the *size* of the area, it generally gives some indication of the overall tendency of the evaluation. A maximal voice, for example, will be represented as only a small shape in the centre of the figure, whereas a voice which is minimal for many of the premises will have a much larger area covered. However, since the size of the area bounded by the line between the points also depends on the arrangement of the axes, one cannot use size very precisely as an indicator of overall tendency. The area will for example be much larger if there are evaluations towards the minimal for several axes next to each other than if every other evaluation is minimal and then maximal, the latter giving a more edgy shape with a much smaller area. Therefore, size can be reliable only when the shapes are relatively similar, and this is most straightforward with shapes approaching a circular shape having the meeting point of the axes at the centre. Nevertheless, shape and size together can still be useful in discovering similarities in the evaluation of different phrases, since graphical shapes are very easy and fast to compare visually. Hence, there are both strengths and weaknesses with this form of representation, where the arbitrariness of the placement and order of the axes is perhaps the clearest weakness, and the 'frozen' snapshot it creates can be seen both as strength and weakness.

[14]For details of the analysis, see Bergsland 2010: chapter 12.
[15]The representations are based on an evaluation of my own experience of the excerpt.
[16]Acousmographe is developed at INA-GRM (http://www.inagrm.com/accueil/outils/acousmographe). Readers that seek a detailed discussion of these evaluations can find these in chapter 12 of my thesis (Bergsland 2010).

https://doi.org/10.1017/S1355771813000137 Published online by Cambridge University Press

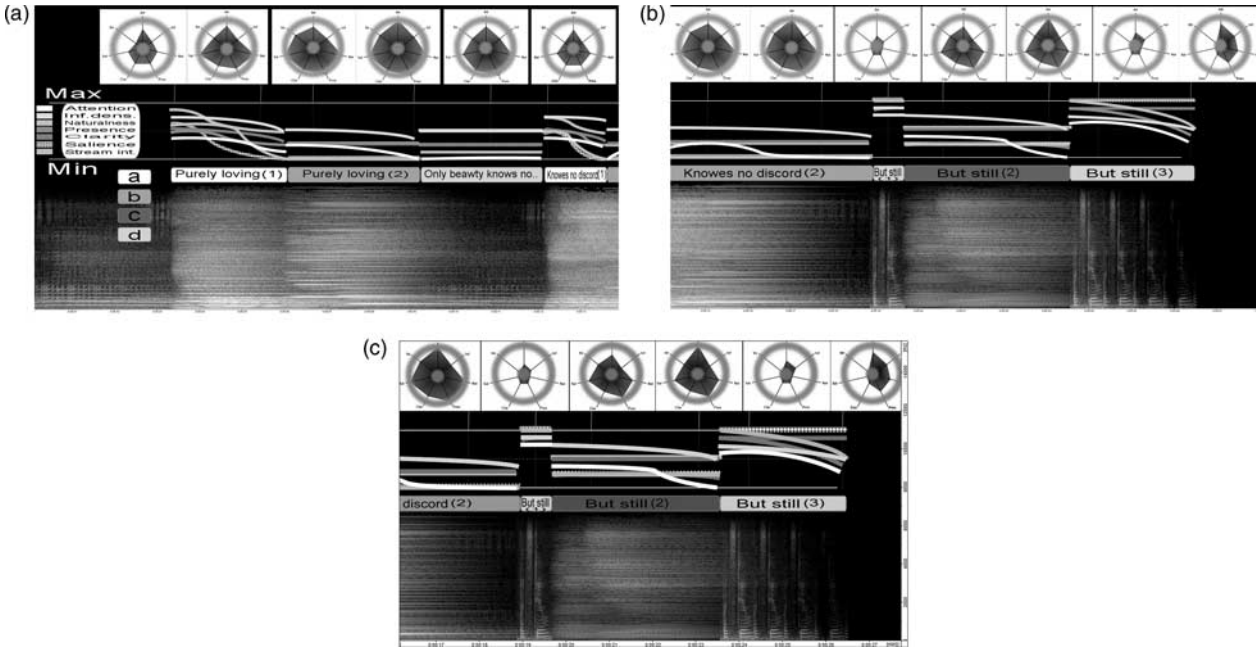

Figure 4. Acousmography in two panels (a, b) of the first 27 seconds of the third movement of Paul Lansky's *Six Fantasies on a Poem by Thomas Campion* entitled *her reflection*. The upper section of each panel shows the axial representations at 13 different temporal locations, while the section just below it shows the time-varying representations of each of the premises in the maximal–minimal model (nomenclature on the left side of figure 4a). The rounded rectangles in the central part of the panels present the text and grouping of each of the phrases in the excerpt.

6.2. Time-varying representations

Below the axial representations in figure 4a and b, one can see the *time-varying* representations displaying the temporal variations of the evaluations of each of the premises in the form of horizontally oriented curves, where each premise has its unique pattern or shade of grey. The curves or lines are placed in a system consisting of an upper and a lower vertical limit representing the maximal and the minimal evaluations.[17] To ease comparison among evaluations, I have drawn the intermediate position between the two extremes as a dotted line. Moreover, I want to make it clear that the placement of these lines within the same 'system', so to speak, is done purely for the sake of having the possibility of viewing the evaluations together: the evaluations are based upon different grounds, so that a similar placement along the continuum does not imply the same thing for different evaluations. Thus, the only thing that these evaluations have in common is that they can all be placed on a continuum between the minimal and the maximal, where the maximal end represents a convergence of the premises as discussed in section 5 above. This does not mean, however, that it is impossible to compare the temporal evaluation of different premises; whether two or more premises have a common tendency or not for a certain phrase, whether they 'peak' at the same time, or whether they reach their bottom levels simultaneously, can still be interesting to note.

While this type of representation remedies the lack of temporal development of the axial representation, the simultaneous representation of seven different axes with seven superimposed lines with different patterns or shades of grey, can be difficult to read and give a slightly 'messy' impression, especially since lines can be hidden behind each other. This can partly be circumvented by viewing them directly in the Acousmographe, where it is possible to turn the visibility of each of the lines on and off. However, in using both representations together one can achieve a fairly good compromise between detail and overview, so that it is possible to get a good idea of the evaluations both temporally and at specific points in a piece.

In this excerpt, for instance, we can see that there is a great deal of variation from phrase to phrase, and that some phrases are indeed moving towards the extreme poles of the continuum. Here, the first 'but still' phrase is very close to the maximal, and the phrases that belong to phrase type b, 'purely loving' and 'knowes no discord', are quite close to the minimal. We can also see that the temporal evaluation of many of the premises in several phrases descends towards the end of the phrases. In this case, this reflects that the ending of many of the phrases finishes with

[17]These lines should perhaps have been made diffuse so as to avoid the impression that they represent a definite, absolute and discrete value, but, due to limitations in the *Acousmographe* software, such a representation was not possible. The reader should therefore keep this in mind when viewing the representation.

https://doi.org/10.1017/S1355771813000137 Published online by Cambridge University Press

decaying comb filter resonances that retain little vocal quality, or that repeating echoes causes an exhaustion of the referential qualities of the vocal phrase – a kind of *sillon fermé* effect, drawing my attention towards sound qualities rather than focusing on the *what*, the *how* and the *who* aspects of the utterance. However, the scope of this article prevents me from going into more detail on the analysis of this excerpt.[18]

7. CONCLUSION

All in all, the maximal–minimal model as a part of the larger framework described in my thesis (Bergsland 2010) provides a systematic way of assessing aspects of the listening experience *beyond* the qualities of the sound itself, hence expanding the more established Schaefferian approaches, while retaining the emphasis on close listening central to these approaches.

The model comprises many ideas proposed by theoreticians dealing with voice in electroacoustic music as well as radio and literature – only in a more systematic manner and including more aspects/dimensions than many of them. Still, I regard it as being far from comprehensive, and open for additions or re-structuring. Among other things, the model could include aspects dealing with multiple voices in social interaction as well as larger temporal structures than the vocal phrase.

While more systematic and formalised frameworks for dealing with listening experiences such as this can always risk being inhibiting and restrictive compared to a freer and more narrative exploration as, for example, demonstrated by Norman (see, e.g., Norman 2000), it also has several advantages.[19] One of the most important is that it provides grounds for *comparisons* of different phrases in a single piece as well segments from different pieces. This can be done with or without the use of graphical visualisations in the form of *axial* and/or *time-varying* representations shown above. Such graphical representations clearly have both an epistemological and a pedagogical value: they create overviews that can allow us to (over)see many aspects at once and note interrelations between them. With the Acousmographe software, this can even be done while listening to the music. Another advantage that I see is that because it deals with a multiplicity of premises (as well as their respective *factors* that make up the criteria for evaluation; see Bergsland 2010: chapters 5–11) it can provide a sort of checklist for aspects that can be relevant for electroacoustic works with voice. This can make composers, performers, students and teachers working with electroacoustic music and voice more conscious about the potential range of meanings and effects that electronically mediated and processed voices can convey.[20] However, the model reaches its fullest potential only when it is seen in relation to the framework as a whole – that is, considering the *experiential domains* (Bergsland 2010: chapter 3), and when one uses it in a process of close and repeated listening as a basis for making evaluations and comparisons.

In many ways, using the maximal–minimal model and the framework as a whole a listener can be made to pay notice to a great many aspects of the experience, and this might subsequently be taken as a basis for a more comprehensive interpretation, hopefully deepening one's understanding of both the piece and the active role that one takes on in the process of listening and making sense of it.

Acknowledgements

This research has been funded by a three-year fellowship from the Faculty of Humanities of the Norwegian University of Science and Technology (NTNU). Carl Haakon Waadeland and Rolf Inge Godøy have been supervisors.

REFERENCES

Aoki, N. and Ifukube, T. 1996. Two 1/f Fluctuations in Sustained Phonation and their Roles on Naturalness of Synthetic Voice. *Proceedings of the Third IEEE International Conference on Electronics, Circuits, and Systems, ICECS '96*. Rhodes, Greece.

Bergsland, A. 2006. From Maximal to Minimal Voice: Concepts for Evaluating Vocal Sounds in Electroacoustic Music. *Electroacoustic Music Studies Network Conference, EMS05* Montreal, Canada.

Bergsland, A. 2010. *Experiencing Voices in Electroacoustic Music*. PhD thesis, Department of Music. Trondheim: NTNU.

Birnbaum, D., Fiebrink, R., Malloch, J. and Wanderley, M.M. 2005. Towards a Dimension Space for Musical Devices. *Proceedings of the 2005 Conference on New Interfaces for Musical Expression*. Vancouver: National University of Singapore.

Bossis, B. 2005. *La voix et la machine: La vocalité artificielle dans la musique contemporaine*. Rennes: Presses Universitaires de Rennes.

Børset, B. 2006. *Støy og stemmer: radioteknologi og stemme i Nathalie Sarrautes pièces radiophoniques*. Det historisk-filosofiske fakultet. Trondheim: NTNU.

Bregman, A.S. 1990. *Auditory Scene Analysis*. Cambridge, MA: The MIT Press.

Chion, M. 1988. Du son à la chose: Hypothèses sur l'objet sonore. *Analyse musicale* **11**: 52–8.

Dyson, F. 1994. *The Genealogy of the Radio Voice*. In D. Augaitis and D. Lander (eds.) *Radio Rethink: Art, Sound and Transmission*. Banff, AB: Walter Phillips Gallery.

[18]See Bergsland 2010: 339–51 for a detailed analysis of this particular phrase.

[19]As Thoresen points out, when setting up pre-existing categories or systematic criteria as guidelines for the listening process, one can risk that listeners lose some of the openness of a less constrained listening process (Thoresen 2007: 132).

[20]Naturally, live performers will also have to take their presence on stage into account (see Emmerson 2007: chapter 4).

https://doi.org/10.1017/S1355771813000137 Published online by Cambridge University Press

Eco, U. 1989. *The Open Work*. Cambridge, MA: Harvard University Press.

Emmerson, S. 1986. The Relation of Language to Materials. In S. Emmerson (ed.) *The Language of Electroacoustic Music*. London: Macmillan.

Emmerson, S. 2007. *Living Electronic Music*. Aldershot: Ashgate.

Ferrara, L. 1984. Phenomenology as a Tool for Musical Analysis. *The Musical Quarterly* **70**(3): 355–73.

Godøy, R.I. 1997. *Formalization and Epistemology*. Oslo: Universitetsforlaget.

Halford, G.S., Wilson, W.H. and Phillips, S. 1998. Processing Capacity Defined by Relational Complexity: Implications for Comparative, Developmental, and Cognitive Psychology. *Behavioral and Brain Sciences* **21**(6): 803–31.

Lakoff, G. 1987. *Women, Fire, and Dangerous Things: What Categories Reveal about the Mind*. Chicago, IL: University of Chicago Press.

Lakoff, G. and Johnson, M. 1980. *Metaphors We Live By*. London: The University of Chicago Press.

Landy, L. 1993. Sound Transformations in Electroacoustic Music. *Composers Desktop Project Quarterly*.

Norman, K. 2000. Stepping Outside for a Moment: Narrative Space in Two Works for Sound Alone. In S. Emmerson (ed.) *Music, Electronic Media and Culture*. Aldershot: Ashgate.

Paas, F., Renkl, A. and Sweller, J. 2004. Cognitive Load Theory: Instructional Implications of the Interaction between Information Structures and Cognitive Architecture. *Instructional Science* **32**(1): 1–8.

Rosch, E. 1975. Cognitive Representations of Semantic Categories. *Journal of Experimental Psychology*. *General* **104**(3): 192–33.

Rosch, E. 1978. Principles of Categorization. In E. Rosch and B.B. Lloyd (eds.) *Cognition and Categorization*. Hillsdale, NJ: John Wiley & Sons.

Rosch, E. and Mervis, C.B. 1975. Family Resemblances: Studies in the Internal Structure of Categories. *Cognitive Psychology* **7**(4): 573–605.

Segnini, R. and Ruviaro, B. 2005. Analysis of Electroacoustic Works with Music and Language Intersections. *ICMC 2005 Proceedings*. Barcelona.

Smalley, D. 1993. Defining Transformations. *Interface* **22**: 279–300.

ten Hoopen, C. 1992. *Abstract and Mimetic Qualities in Electroacoustic Music: Technology Amsterdam*. Amsterdam: Rodopi.

Thoresen, L. 2007. Spectromorphological Analysis of Sound Objects: An Adaptation of Pierre Schaeffer's Typomorphology. *Organised Sound* **12**(2): 129–41.

Verfaille, V., Gustavino, C. and Depalle, P. 2005. Perceptual Evaluation of Vibrato Models. *Proceedings of Conference on Interdisciplinary Musicology, CIM05*. Montreal, Canada.

Wesling, D. and Slawek, T. 1995. *Literary Voice: The Calling of Jonah*. Albany: State University of New York Press.

Wishart, T. 1996. *On Sonic Art: A New and Revised Edition*. Ed. S. Emmerson. London: Routledge.

Young, J. 1996. Imagening the Source: The Interplay of Realism and Abstraction in Electroacoustic Music. *Contemporary Music Review* **15**(1): 73–93.

RECORDINGS CITED

Harvey, J. 1990. *Mortuos Plango, Vivos Voco* (1980). On Various Artists, *Computer Music Currents 5: Music with Computers*. Mainz: Wergo WER2025-50.

Lansky, P. 1994. *Six Fantasies on a Poem by Thomas Campion* (1979). On *Fantasies and Tableaux*. Composers Recordings, CRI 683.

Ligeti, G. 2006. *Nouvelles aventures* (1965). On *Requiem; Aventures; Nouvelles aventures*. Wergo WER 6925 2.

Parmerud, Å. 1994. *Les objects obscures* (1991). On *Osynlig Musik/Invisible Music/Musique invisible*. Phono Suecia PSCD 72.

Schwitters, K. 1992. *Ursonate: Sonate In Urlauten* (1932). On *Ursonate*. Hat Hut Records ART CD 6109.

Stockhausen, K. 2001. *Gesang der Jünglinge* (1956). On *Elektronische Musik 1952–1960*. Stockhausen-Verlag Stockhausen 3.

Viñao, A. 1994. *Chant d'ailleurs* (1992) On *Hildegard's Dream*. MUSIDISC MU 244942.

Wishart, T. 1992. *Red Bird* (1977). On *Red Bird: Anticredos*. October Music Oct 001.