

## Sju ... Who?

John Young  
Music, Technology and Innovation Research Group  
De Montfort University  
Leicester LE1 9BH, United Kingdom  
jyoung@dmu.ac.uk

### ABSTRACT

This paper presents a discussion of the motivations and techniques used in the author's electroacoustic work *Sju*. The piece is based on digital recordings of a Swedish word that, in daily Swedish, is pronounced in two distinctive ways, thereby presenting a real-life sonic opposition ready for musical exploration. This is related to the concept of sonological competence (drawn from soundscape studies) since in both versions the word is awkward to pronounce for non-Swedish speakers, an issue stemming directly from the composer's first-hand experience with Swedish. The approach used in the manipulation of these sounds is outlined, especially the development of material around the variations in noise attack which define the different 'versions' of the word, and the extrapolation of the play between noise and speech into the larger formal structure of a work.

**KEYWORDS** Electroacoustic composition, sonological competence, vocal sound, Swedish language, granular synthesis.

### 1. SOURCES

*Sju* (1999) is an electroacoustic work created with material recorded while I was working in the studios of Elektroakustisk Musik i Sverige (EMS), with the assistance of a Swedish Institute guest scholarship. The piece is based on the variable pronunciation of the Swedish word 'sju', the meaning of which is quite prosaic—the number seven—but to my ear the sounds were immensely interesting. There are two very distinct pronunciations of 'sju', one is strictly correct 'pure' Swedish, and another a more colloquial version, to the extent that a speaker's use of one or other of these pronunciations could be taken as an indication of social class (though I discovered in recording different versions of it that there are also some subtle in-between variants)<sup>1</sup>. When I first arrived in Sweden two of the staff of EMS used this 'pure' version of this word as a fun Swedish competency test for me since there is no truly equivalent sound in English. I did not do so well in that test (!), but the elusive nature of the sound, plus the realisation that there are two quite contradictory sound shapes unified by a common semantic meaning suggested the possibility of an electroacoustic piece to me. The difference in pronunciation is found in the initial brief noise

---

<sup>1</sup> One of the Swedes involved in the recording of material for the work told me that she was teased at school for sounding too 'posh' with her pure version of the word. There are also many other examples in Swedish of similar kinds of formal and informal variants on pronunciation, some of which have resulted in changes to accepted spelling.

component of the word, and I recorded the different versions with the assistance of a number of Swedish people. Five of these are grouped in this sound example, and the sonogram is displayed in fig. 1. The defining polarities of the 'sju' phenomenon are heard in the distinction between the 'shh' and 'whh' attack transients, most clearly typified by the first and last in this sequence, though several versions are bracketed here to indicate the variance within that basic distinction.

Ex. 1

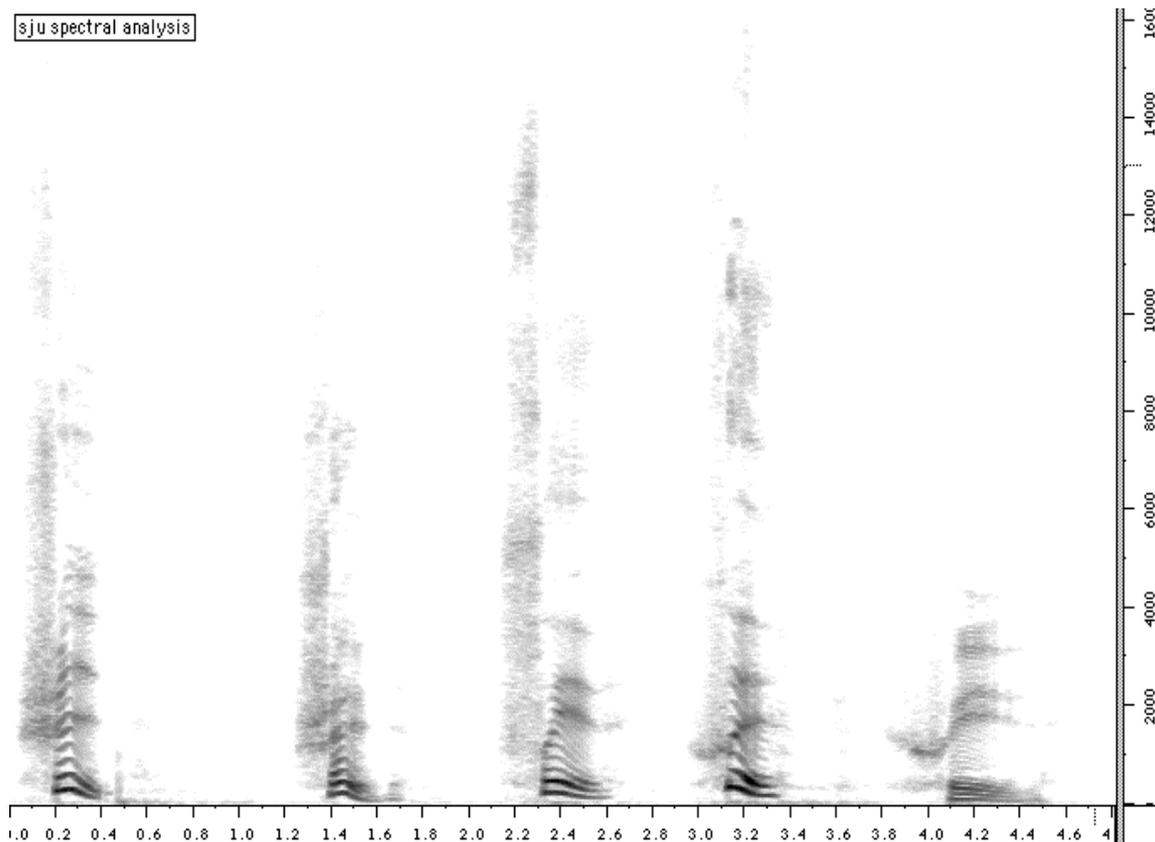


Fig. 1 Spectrogram of five versions of the Swedish word 'Sju'.

Early on I had the idea that I may want to contextualise whatever electroacoustic extensions of the sound that might become possible with a sense of personalities recurring or clustered through the work, and to that end I ensured that I had a range of voice types, genders and ages. In the recording sessions I asked each of the participants (most of whom were musicians) to speak the word in their own natural version, and a number of them became absorbed in repetition of the word, which also extended to some improvised play on it, one instance of which is retained in the work (at 1'28"). One subject, a young boy who, when invited to play with the word, produced a startling cadenza of speech-song (discussed later). Overall, repetitions of this kind provided me with a large number of minute sonic variations for the word as produced by each speaker, underlining a pre-

compositional strategy which can assist in giving suppleness and realism to repetition in acousmatic music.

## **2. LANGUAGE AND SOUND ISSUES**

The use of vocal sound sources in electroacoustic music has obvious attractions: the enormous range of sound types that can be produced and subtle inflections within them, the potential to embrace language while developing new morphological constructs, and the obvious potential to develop ideas around the 'humanity' of the sounds themselves. But the voice also has a useful potential to emulate non-vocal complex sound shapes since noise, pitch and resonance can be extensively and expressively modulated. In short, human vocal sound presents a self-contained continuum of representation and abstraction circumscribed by the potential for language and the extreme flexibility of the sound-producing mechanism itself. Vocal sound production, even the semblance of it, invites us to identify with and infer meaning from it because of our reliance on linguistic and paralinguistic utterance as a form of communication, as well as our conditioned knowledge of involuntary utterance and the timbre and gesture spaces defined by physically-defined factors such as vowel formants and breathing patterns. With 'sju' I encountered a clear distinction between sounds whose intrinsic spectral complexity made them worthy of development in a composition, but which also drew me into a desire to 'comprehend' it better via the compositional process following my attempts to accurately enunciate it.

As a result, this became for me an issue of *sonological competence*, a concept from the writing of Otto Laske integrated into soundscape studies by R. Murray Schafer. Schafer defines sonological competence as the facility which 'unites impression with cognition and makes it possible to formulate and express sonic perceptions' (Schafer 1994: 274). In other words, it is a measure of the bridge between the reception of sound, comprehension of its structure and the mechanics of its formation as expressed by, for instance, the ability to repeat or mimic that sound. This shifted to an issue of 'performance', since I was also encountering a phonemic construct outside of my linguistic experience.<sup>2</sup> Denis Smalley suggests our perception of the proprioceptive foundation to sounds derived from human agency is part of our complete psychological response to hearing. He proposes that we follow and interpret a sound's energy and spectral shaping through time with the potential to correlate this with muscular actions, tensions and releases (Smalley 1997: 111-112). This was certainly relevant to my interest in the 'sju' phenomenon since, in the games I had to play trying to articulate it, I had difficulty matching what I thought were the correct oral and vocal patterns with a truly satisfactory sonic result.

In another sense, the idea of sonological competence offers one perspective on a question in electroacoustic composition which is frequently posed informally, but seldom satisfactorily answered: what is an 'interesting' sound? Apart from

---

<sup>2</sup> The shift from 'competence' to 'performance' is discussed in musical terms by Truax (1994: 49-50).

being drawn to the way unvoiced noise is modulated across the versions of the word, its melding into a voiced pitch and the contradiction inherent in its pronunciation by the Swedish people, I was prompted to explore the qualities of these sounds given the linguistic challenge it represented. That is to say: this sound, beautiful in its fusion of noise and pitch, seemed to hold an unattainable sonic essence, and attempts to reproduce it simply produced even more variants of it.

### **3. SOUND PROCESSING IN *SJU***

The notion of sonological competence was a formative factor in the approach taken to the conception and development of this piece, encouraging me to retain a certain amount of source and contextual coherence throughout. I was aiming to encourage a listener to gain some sense of my confrontation with the pronunciation of 'sju' by virtue of the many variants used and their electroacoustic extensions (a process which might be thought of as a form of pragmatic analysis), as well as tying this to the complicating factor of the conflicting nature of the pronunciations as they occur in reality. To a certain extent that influenced the way processing techniques were employed, and the range of sounds drawn on in the wider process of developing and mixing materials. Initially, for instance, I thought of the possibility of elaborating on the 'tension' of the noise attack through the pasting and overlapping of these into an unstable and skittish 'enunciation' (though this was not used in the piece).

Ex. 2

Because the main distinguishing feature of these sounds is found in the noise attack to the word (typically in the order of 150 milliseconds), I looked for a way to extend these in time, so that the spectral variation could be more fully appreciated and so that the more extended sound shapes could contribute to longer and overlapping phrases. From the perspective of sonological competence this was also a way of expressing the 'sju' phenomenon at a magnified and more comprehensible rate of evolution than would be the case in normal speech rhythm (especially for an uninitiated listener). To this end, I 'stretched' edited clips of transient using closely packed bands of low frequency comb filters fed into a doppler effect algorithm rotating at just above audio frequency, causing the noise output of the comb filters to 'resonate' spatially, but elongating the distinctive colour of each transient. In particular, variations in doppler rotation frequency introduced added textural vivacity to the resulting sound shapes. In these two sound examples, the 'dry' input and then 'processed' output are heard, the latter appearing from 0'46" in the completed work.

Ex. 3 and Ex. 4

The production of these particular sounds was a pivotal point in the composition process, since they retain a clear generative relationship with the original 'sju' transients, but now have a continuous morphological identity in their own right, perceptually akin to a 'fluttersong'. This 'flutter' sound could therefore be the

base of a new set of transformations with a familial relationship derived from the new morphology. In these examples, frequency shifting is applied such that vocal formants are lost, weakening the link with the original vocal transient, producing a fresh group of sounds that are cognates of the 'flutter' morphology.

Ex. 5

The construction of a phrase that presents these cognates in a continuous stream may give the impression of sound-play constrained by the physical/behavioural properties of the sounds or their imagined source. This example, though not actually included in the final work, is given to illustrate something of the way my attempts to manipulate these materials evolved.

Ex. 6

I agree with Trevor Wishart's assertion that it is the perceived rather than generative relationships amongst sounds that most significantly determine the way we tend to construct musical relationships, since the composer's knowledge of generative relationships does not necessarily translate into audible connections between sounds (Wishart 2000: 22-3). However, the contextual associations created in composition may enable groups of cognates to intersect and develop in parallel while evidence of their generative relationship is signalled. Two instances of this kind of structural intersection in *Sju* are at 0'45" and 1'34" - 1'39". In the first instance the spoken voice and flutter transformation articulate the same transient, while in the second example, the spoken and flutter versions are in contradiction.

Ex. 7

One specific influence of the sonological competence challenge was the idea of creating a rapidly shifting palette of sound derived from the noise colours themselves. This was influenced by the assumption that a difficult-to-pronounce word is even more difficult when repeated, or tossed off in the course of a sentence.<sup>3</sup> In order to electroacoustically model an imaginary version of that condition, I used an asynchronous granulation technique, whereby a continuous file comprised of many different versions of the 'sju' transient spliced together was created, and a new sound file constructed by sampling and reassembling chunks out of that. Texturally this 'sampling within the sample' creates rhythmic play around the different noise colours that form a cluster of sonic cognates of related provenance, but with varying outward contour. This attempts to articulate a 'variation space' for the sound, in itself an analogy and extension of the linguistic dilemma of the 'sju' phenomenon.

Ex. 8

Because long 'grains' are used here (up to about 350 milliseconds) I regard this as a form of randomised mixing/editing rather than 'granular synthesis' and the length of grain was influenced by the motivation to maintain recognition of the

---

<sup>3</sup> The 'sju' sound is also the subject of a tongue twister familiar to all Swedes, and is the focus of a new work in progress.

variance amongst the noise attacks. But as 'mixing', this approach also was catalyst for the interleaving of the original vocal sounds with processed versions. The generating morphology of the granular process is, in effect, separate from the sounds themselves (defined by the limits and density of grain length). Thus the textural patterns created by this process can project into wider forms of musical development when transformational variants are injected into the mixing/editing process (in this case the 'flutter' sounds previously described). Some of the clearest examples of material developed in this way are found in the work at 4'12"-4'52" and 5'05"-6'11".

Ex. 9

This concentration of different noise spectra led to mixing these with other sounds that had relatively superficial relationship with the vocal sounds at the purely sonic, rather than the contextual or even transformational level. In this example, the 'sju' attack transients are mixed with samples from ice skaters recorded in Stockholm's Kungsträdgården (from 6'12"), again mixed and edited using asynchronous grains of up to 350ms. The sonic 'links' which guided me in this case were the dynamic surges and the spectral shaping of the ice skating sounds (which can easily be imitated through shifts in vowel formant), and this is anticipated at 6'00" in the lengthening of the 'flutter' envelopes at that point.

Ex. 10

Another instance of this kind of sonic free association is in the use of a sample of whispering resonated (physically) through a flute, for example at 3'07". Here the flute/voice sound is used to add spectral complexity to a granular mix similar to the example above, by imposing the granular mix's amplitude envelope onto that of the voice/flute. This process of allowing a dynamic pattern to function as a transferable rhythmic identity through a structure is one I have used extensively in acousmatic composition.

Ex. 11

One of the consequences of intensive studio listening is a heightened sensitivity to the character and subtleties of materials constantly being re-auditioned, which can result in a tendency to begin hearing sounds developed in the studio 'inside' sounds encountered in day-to-day experience, and vice versa. In making the sorts of sound links outlined above I was to a large extent attempting to objectify sound associations of that kind.

A synchronous granulation variant of 'sampling within the sample' was also used in *Sju*, with comparatively long grain lengths, reconstructed at low density, producing a reconstituted sound file that resembles 'snapshots' of the file's evolution in time. It was used to process the entire word 'sju', as performed by the youngest of my participants in the recording sessions, a small boy named Fivos. Fivos's extraordinary extemporisation on the word was textured using a gradual increase in 'grain' density. The morphology of the transformation is such that the word is disjointed to begin with, becomes more coherent, and then

overlaps to produce a tight delay/phase effect. The key idea for me here is that of coherence: there is level of continuity within the vocal sound that a listener (drawing on their general experience of vocal sound shapes) might vicariously interpret as 'whole' or 'truncated' (in *Sju*, this would also hopefully be confirmed by context, since the word has already been presented in its entirety several times). In this example, that aspect of contextual meaning gives the dynamic shape of the transformational morphology an additional significance beyond that of simply an 'accelerating rhythm approaching iteration'. In the context of the work (2'44") this acceleration heralds a new version of the 'flutter' morphology, emphasising the gathering of momentum.

Ex. 12

I use what are essentially algorithmic procedures like this to construct new sound formations for which I could not necessarily predict the outcome. However, because the algorithmic process is not a sonic entity, nor something legitimating the design of the music itself, I always apply a critical ear to the results, modifying these where a musical intention makes it necessary (either through tweaking the processing parameters or through the editing or reshaping of some specific feature of the sounds). By a musical intention I mean factors such as the nature of the rhythmic shaping of the material (especially with a view to its potential role as an element in the dynamic evolution of the music's flow) and the extent to which source recognition of a sound might be retained or recontextualised. In the final event I feel my primary relationship to be with my *sound* materials, to the extent that the technological methods used to construct a new sound become of secondary importance to the actual sonic qualities themselves. In working with natural sound samples drawn from real-world experiences and situations, the potential for the sound source to carry meaning—contextual, metaphorical, linguistic—remains very central to the way I tend to direct and explore the potentials of sound processing. In reality, of course, it is the interaction of processing artefacts and the intrinsic features of the sounds themselves that fully determine the evolution of the materials. It is from this perspective that I regard the relationship between signal processing method and sound file as not existing in arbitrary relationships but ones where the intrinsic structure of sound and the shaping tendencies of the processing structures must relate to articulate an essential idea or view of a sound.

#### **4. STRUCTURE AND CONTEXT**

The most salient sounding feature common to all the versions of the word 'sju' is its subdivision into consonant and vowel (unvoiced and voiced) components, and the overriding musical context created from the 'sju' sound is a dislocation of these two elements. In developing the overall structure of the piece this basic distinction between consonant and vowel parts of the word was kept in mind as a general polarity of noise and pitch. The noise aspect provided material for the main audibly 'developmental' part of the piece such as the passage 3'00" - 3'35", which presents a comparatively saturated clustering of noise shapes, progressively dissipating to a solo 'sju' transient. The motion of noise into pitch as

a generative model was also a consideration, for example at 1'39" - 2'08" where the mixing of many noise-bands might be thought of as a highly embellished version of the noise-pitch transition. Pitched resonance provides more long-range underpinning of texture, and also articulates phrase structures in conjunction noise-based rhythmic and textural play, for example at 4'48".

Ex. 13

At the end of the work, the vowel-like resonances and attack transients are reintegrated to some extent. I found that 'flutter' noise attacks developed from one person could be edited alongside the purely vowel component of the word as spoken by another, allowing the word 'sju' to be resynthesised from two processed components (6'56"-7'06"). In fact this kind of mix and match reconstruction is quite plausible even with unprocessed components.

Ex. 14

Throughout *Sju* the original context is reinforced through the statements of the complete word, which are constructed to highlight the contradictory pronunciations, as well as some of the sense of 'play' with the sound of the word that was captured in recording. Examples are found in the passages at 39"-45", 1'21"-1'32", 2'44"-2'53", 4'05"-4'12" and 6'56"-end.

## 5. CONCLUSION

While *Sju* attempts to evoke and musically amplify aspects of a very specific linguistic encounter, it also underlines fundamental aspects of the innovative potentials of the electroacoustic medium itself. Since sound recording has provided us with the resources to capture real-world sounds intact, study and manipulate them in a state suspension out of real-time (or, perhaps more significantly, out-of 'performance') we have a means of investigating, through sound, something of the substance of our experience in a very direct way—the ways we listen as well as the ways we communicate. The sheer act of turning the microphone towards cultural and environmental phenomena can provide the materials which, by virtue of their permanence as sound documents, may encourage ever more intense and questioning listening—both in and out of the studio. While the many facets of computer-enabled transformation are fundamental to the development of digital sound art, the groundwork formed by natural sounds themselves and the networks of representation and association we derive from them must constitute one of the strongest forces underlying electroacoustic music's potential as vibrant and meaningful sonic art.

*Sju* is dedicated to everyone at EMS Stockholm, with special thanks to Inger, Göran, Paulina and her composition class, Perikles and Fivos.

## REFERENCES

Schafer, R. Murray 1994. *The Soundscape: The Sonic Environment and the Tuning of the World*. Rochester, Vermont: Destiny.

- Smalley, Denis 1997. Spectromorphology: Explaining Sound Shapes. *Organised Sound*, 2:2, 97, 107-126.
- Truax, Barry 1994. *Acoustic Communication*. Norwood, NJ: Ablex.
- Wishart, Trevor 2000. Sonic Composition in Tongues of Fire. *Computer Music Journal* 25:2 S00, 22-30.
- Young, J. 2000. *Sju* (1999) In *New Zealand Sonic Art 2000* [CD recording]. Hamilton: Music Department, University of Waikato.