

41st International Computer Music Conference

ICMC 2015

[icmc2015.unt.edu](http://icmc2015.unt.edu)



Looking Back, Looking Forward

Denton, September 25 - October 1

Proceedings

Hosted by the Center for Experimental Music and Intermedia

**UNT** UNIVERSITY  
OF NORTH TEXAS®

**ICMC 2015**  
icmc2015.unt.edu

# **Looking Back, Looking Forward**

Denton, September 25 – October 1

## **Proceedings**

Hosted by the Center for Experimental Music and Intermedia

**UNT – University of North Texas**

Proceedings of the 41st International Computer Music Conference 2015  
September 25 — October 1, 2015  
University of North Texas, Denton, TX, USA

Looking Back, Looking Forward

Richard Dudas, editor

ISBN-10: 0-9845274-4-3

ISBN-13: 978-0-9845274-4-1

Copyright © 2015 - All copyright remains with the individual authors

Published by:

The International Computer Music Association  
2040 Polk Street, Suite 330  
San Francisco, CA 94019  
United States of America



## ICMC 2015 Paper Selection Committee

Miriam Akkermann, *Berlin University of the Arts*, Germany  
Jamie Bullock, *Birmingham Conservatoire*, United Kingdom  
Nicolas Castagné, *Univ. Grenoble Alpes, INP, ICA laboratory & ACROE*, France  
Paul Doornbusch, *Australian College of the Arts*, Australia  
Carl Faia, *Brunel University, London*, United Kingdom  
Pete Furniss, *University of Edinburgh*, United Kingdom  
Rob Hamilton, *Stanford University*, United States  
Mara Helmuth, *University of Cincinnati*, United States  
Dobromiła Jaskot, Poland  
Steven Kemper, *Rutgers, The State University of New Jersey*, United States  
David Kim-Boyle, *University of Sydney*, Australia  
John MacCallum, *CNMAT, University of California, Berkeley*, United States  
David Medine, *University of California, San Diego*, United States  
Miller Puckette, *University of California, San Diego*, United States  
Mihir Sarkar, *MIT Media Lab*, United States  
Alexander Sigman, *Yamanashi Gakuin University*, Japan  
Tamara Smyth, *University of California, San Diego*, United States  
Marij van Gorkom, *De Montfort University*, United Kingdom  
Lonce Wyse, *National University of Singapore*, Singapore

## ICMC 2015 Music Selection Committee

James Andean, *Sibelius Academy, University of the Arts, Helsinki*, Finland  
Françoise Barrière, *MISAME*, France  
Gonzalo Biffarella, *Universidad Nacional de Córdoba*, Argentina  
Elizabeth Hoffman, *New York University*, United States  
Christopher Hopkins, *Iowa State University*, United States  
Vera Ivanova, *Chapman University, ACF LA, Synchromy*, United States  
Konstantinos Karathanasis, *University of Oklahoma*, United States  
David Kim-Boyle, *University of Sydney*, Australia  
Adrian Moore, *The University of Sheffield*, United Kingdom  
Tae Hong Park, *New York University*, United States  
Jøran Rudi, *NOTAM*, Norway  
Andrea Szigetvári, *Hungarian Computer Music Foundation*, Hungary

## ICMC 2015 Paper Review Committee

Miriam Akkermann  
Alessandro Altavilla  
Torsten Anders  
Ted Apel  
Mark Ballora  
Leah Barclay  
Edgar Berdahl  
Fearn Bishop  
Andrew Brown  
Jamie Bullock  
Christopher Burns  
Juan José Burred  
Baptiste Caramiaux  
Chris Chafe  
Marko Ciciliani  
Miha Ciglar  
Alessandro Cipriani  
David Coll  
Arshia Cont  
David Cope  
Nuno N. Correia  
Cathy Cox  
Roger Dannenberg  
Giovanni De Poli  
Marco Donnarumma  
Paul Doornbusch  
José Echeveste  
Aaron Einbond  
Simon Emmerson  
Tom Erbe  
Cumhur Erkut  
Georg Essl  
Carl Faia  
Rebecca Fiebrink  
Rajmil Fischman  
Dominique Fober

Ivan Franco  
Paul Fretwell  
Jean-Louis Giavitto  
Maurizio Giri  
Rama Gottfried  
Owen Green  
Michael Gurevich  
Rob Hamilton  
Alex Harker  
Ian Hattwick  
Christopher Haworth  
Lauren Hayes  
Jason Hockman  
Avrum Hollinger  
Stuart James  
Alexander Refsum Jensenius  
Sergi Jorda  
Jean-Marc Jot  
Emmanuel Jourdan  
Orestis Karamanlis  
Gregorio García Karman  
Steven Kemper  
Ji Chul Kim  
David Kim-Boyle  
Michèl Koenders  
Juraj Kojš  
Olivier Lartillot  
Sylvain Le Groux  
Johnathan F. Lee  
PerMagnus Lindborg  
Cort Lippe  
John MacCallum  
Thor Magnusson  
Mikhail Malt  
Cory McKay  
Andrew McPherson  
David Medine

Flo Menezes  
Nicolas Misdariis  
Chikashi Miyama  
Dafna Naphtali  
Peter Nelson  
Jérôme Nika  
Tim Opie  
Miguel Ortiz  
Francois Pachet  
Kevin Parks  
Timothy Place  
Archontis Politis  
Laurent Pottier  
Miller Puckette  
Ville Pulkki  
Curtis Roads  
Ilya Rostovtsev  
Robert Rowe  
Jøran Rudi  
Adriana Sa  
Mihir Sarkar  
Carla Scaletti  
Diemo Schwarz  
Stephen Sinclair  
Mattias Sköld  
Tamara Smyth  
John Sullivan  
Rafael Valle  
Christophe Vergez  
Graham Wakefield  
Marcelo M. Wanderley  
Johnty Wang  
Andreas Weixler  
Lonc Wyse  
Michael Zbyszyński  
David Zicarelli

## **ICMC 2015 Music Review Committee**

James Andean  
Kevin Austin  
Miguel Azguime  
Natasha Barrett  
Françoise Barrière  
Andrew Bentley  
David Berezan  
Nicolas Bernier  
Gonzalo Biffarella  
Benjamin Broening  
Ludger Brümmer  
Rodrigo Cádiz  
Joel Chadabe  
Marek Chołoniewski  
Marko Ciciliani  
Ricardo Climent  
Ricardo Dal Farra  
Antonio Ferreira  
Jason Freeman  
Mara Helmuth  
Elizabeth Hoffman  
Christopher Hopkins  
Vera Ivanova  
Konstantinos Karathanasis  
David Kim-Boyle  
Joe Klein  
Colby Leider  
Elainie Lillios  
Liao Lin-Ni  
PerMagnus Lindborg  
Cort Lippe  
Minjie Lu  
Mario Mary  
Andrew McPherson

Flo Menezes  
Scott Miller  
Andrea Molino  
Adrian Moore  
Valerio Murat  
Vassos Nicolaou  
Erik Nyström  
João Pedro Oliveira  
Felipe Otondo  
Tae Hong Park  
Rui Penha  
Michal Rataj  
Michael Rhoades  
Jøran Rudi  
Federico Schumacher  
Antonio Sousa Dias  
Kurt Stallmann  
Adam Stansbie  
Nikos Stavropoulos  
Pete Stollery  
Martin Supper  
Andrea Szigetvári  
Robert Thompson  
German Toro  
Dan Tramte  
Graeme Truslove  
Yu-Chung Tseng  
Anders Tveit  
Katerina Tzedaki  
Rodney Waschka  
Chapman Welch  
Daniel Weymouth  
Scott Wyatt  
Lidia Zielińska

### **Piece+Paper Concert and Presentations**

---

- 342 *Antony: A Reimagining*  
John MacCallum, Matthew Goodheart, Adrian Freed
- 346 *Notes on "Culture of Fire" for Analog Neural Network Synthesizer, Geiger Muller Counters and Computer*  
Scot Gresham-Lancaster
- 352 *Audio Spray Gun 0.8 – the Generation of Large Sound-Groups and Their Use in Three-Dimensional Spatialisation*  
Richard Garrett
- 356 *materialssoundmusic: a Computer-Aided Data-Driven Composition Environment for the Sonification and Dramatization of Scientific Data Streams*  
Marco Buongiorno Nardelli
- 360 *An Approach to the Generation of Real-Time Notation via Audio Analysis: The Semantics of Redaction*  
Lindsay Vickery

### **Demonstrations III**

---

- 364 *In and Out, Over and Under: An Interactive Audio-Visual Installation Responding to Percy Grainger's Free Music and the Grainger Museum*  
Roger Alsop

### **Poster Session III**

---

- 368 *TosCA: an OSC Communication Plugin for Object-Oriented Spatialization Authoring*  
Thibaut Carpentier
- 372 *The Effects of Pitch and Dynamics on the Emotional Characteristics of Piano Sounds*  
Chuck-jee Chau, Andrew Horner
- 376 *Media Modules: Intermedia Systems in a Pure Functional Paradigm*  
Mark Santolucito, Donya Quick, Paul Hudak
- 380 *Gesture Capture, Processing, and Asynchronous Playback within Web Audio Instruments*  
Benjamin Taylor, Jesse Allison
- 384 *Rasping Music: Remodeling Early Minimalist Music through Mechatronic Sound-Sculpture*  
Mo H. Zareei, Ajay Kapur, Dale A. Carnegie
- 388 *Mapping Tone Helixes to Cylindrical Lattices Using Chiral Angles*  
Hanlin Hu, Brett Park, David Gerhard
- 392 *Mirror Mind: New Possibilities for Overall Interactive Design in New Music-Media Theatre*  
Yi Qin, Da-Lei Fang, Zhi-Bo Xu, Yan Da

## An Approach to the Generation of Real-Time Notation via Audio Analysis: *The Semantics of Redaction*

Lindsay Vickery

Edith Cowan University

[l.vickery@ecu.edu.au](mailto:l.vickery@ecu.edu.au)

### ABSTRACT

*This paper discusses approach mechanically and conceptually to generation notation in real-time in MaxMSP in the work The Semantics of Redaction. The work uses analysis data to generate scrolling, but conventional looking notation. The issues discussed include scaling and mapping of audio data to visual representation, the formal structure of the work and its relation to the generative processes, semantic principles guiding the notational concept and the development of a performance practice for the work. The intention of the work is to provide a flexible vehicle for a performer to explore audio from the media-world in a near-real-time manner using a recent, topical speech recording as the generative artifact for each performance.*

### 1. INTRODUCTION

*The Semantics of Redaction* [2014] (SoR) is a work for solo performer, pre-recorded audio and generative score. To perform the work, a recent, topical speech recording is chosen and loaded. The intent here is that the performer chooses a new recording for each performance emphasizing the “near real-time”, “daily news”-like aspect of the process. The audio is played through the work’s analysis patch and the analysis parameters (frequency, amplitude, brightness, noisiness and attack) are scaled to provide desirable outputs, in terms of range, density and diversity of the generated notation. Instrumentation is chosen by the performer as a commentary upon the subject matter of the recording. It is intended that the generative score is simple enough to perform after a brief period of familiarisation.

The work was written for percussionist Vanessa Tomlinson and officially premiered in her program 8 Hits [1], an ABC Classic FM live broadcast on November 1 2014. Tomlinson chose a speech made by climate-change activist Naomi Klein as part of the promotional tour for her book *This Changes Everything* [2014] [2].

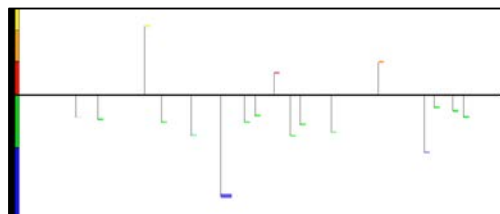
Copyright: © 2015 Lindsay Vickery. This is an open-access article distributed under the terms of the [Creative Commons Attribution License 3.0 Unported](https://creativecommons.org/licenses/by/3.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Along with its sister work *Lyrebird* [3], *SoR* began development during Tomlinson’s residency at the *Orpheus Institute for Advanced Studies & Research In Music* in December 2013.

Tomlinson is both a strong improviser and reader and the work specifically aimed to open a space in which she could interact with a chosen recording with a combination of precision and freedom. Like previous “series” by this composer, *Delicious Ironies* [2001] [4], *Splice* [2002] [5] and *Lyrebird* [2014] *SoR* is intended to be a structured improvisation environment that takes on different qualities in each performance.

### 2. GENERATIVE APPROACH

The notation for *SoR* (see Figure 1) is principally created by using accents detected in a speech recording in real-time to generate graphical symbols of varying vertical position, size and colour, determined by the frequency, amplitude and timbre of the recording at the accent point.



**Figure 1.** The opening section of *The Semantics of Redaction* [2014].

Analysis of the recording is derived from Tristan Jehan’s [6] *analyzer~* object, using attack, frequency, amplitude, brightness and noisiness data. The flow of analysis data is captured at the point of each detected attack. By default, the values defining an attack are an increase in the incoming signal of +10db in a period less than 100ms, however the performer may alter these values to achieve a better score/visualisation result.

The data is mapped in the following manner:

Frequency: Vertical height and notehead colour hue  
Amplitude: Notehead size  
Brightness: Notehead colour saturation  
Noisiness: Notehead colour luminance



The data is drawn to MaxMSP's *jit.lcd* object using the standard *paintrect* command. A stem is also drawn on the left side of the notehead above a central "beam". A left side stem is used regardless of whether the notehead is above or below the beam as an aid to visualising the precise onset of the accent and therefore performer's note execution.

Because of the variation of the acoustic characteristics of potential recordings and performer preferences, an analysis window allows the performer to scale incoming data (See Figure 2). Scaling changes may be necessary because of variations in the average frequency of the speaking voice or ambient noise in the recording. From the performer's perspective the attack rate and sensitivity values create changes in the density of note events in the score.

The analysis window provides the performer with "peak", "trough" and "mean" values for each of the analysis parameters from the recording, as well as a graphical display showing the most recent 100 values for each parameter. A swatch graphically displays the variations in notehead colours. A *preset* object allows the performer to save the chosen scalings for each new recording.

Although it would be possible to automate the scaling process, allowing the performer to set the values provides a means to individualise the score that is generated by each recording: for example to achieve variations in the number of noteheads that fall within different "registers".

Unlike *Lyrebird*, which has a range of colour interpolations and generates a spectrogram-like score, only five hues (Yellow, Orange, Red, Green and Blue) are used in *SoR* (although there is continuous variation in the saturation and luminance of each colour). This approach was taken to achieve a clear differentiation between five registers in the score, which are conceived as instrument groups by the percussionist: for example ranging from yellow (high-pitched instruments) to blue (low-pitched instruments). The notehead size is intended to correlate to dynamic range.

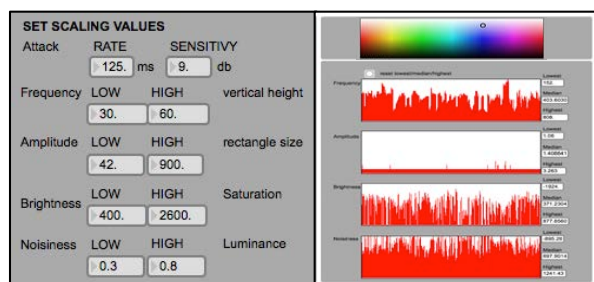


Figure 2. The Analysis window.

In addition to the generative process described above - which occur during the *Opening* and *Body* sections of the work - there are three other modes: *Commentary*, *Interlude* and *Redactio*. These modal changes form sections in the

structure of the work. In each of these sections graphical symbols are added to the notation. Figure 3 gives an indication of the morphologies of the "families" of symbols. There are 76 symbols in total.



Figure 3. Families of Graphical Symbols employed in different sections of *The Semantics of Redaction*.

In the *Commentary* sections the noteheads are extended indicating the note event should be held proportionally for that duration. The wave-like black lines that widen or narrow, indicate timbral or dynamic variation of the note that remains constant, (See Figure 4a.)

In the *Interlude* section some noteheads appear without stems indicating non-accented/slurred gracenotes, rather like liquescent neumes in medieval notation, while stems indicate accents. Graphical symbols in the form of black vertical lines are added and are interpreted as note or noise repetitions (See Figure 4b.). During this section the performer plays solo: the recorded speech is not heard by the audience (although it still drives the generation of notation). Audio from the speech is recorded into buffers that are heard during the "redactio" sections.

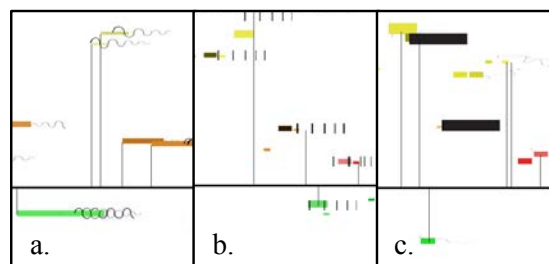


Figure 4. *Commentary* (4a), *Interlude* (4b) and *Redactio* (4c) sections from *The Semantics of Redaction*.

The *Redactio* section (See Figure 4c.) combines elements of the *Body* and *Commentary* sections, however black rectangles (redactangles) accompanied by beating sine tones, obscure elements of the score (and recording). Prerecorded samples from the previous *Interlude* are manipulated (via fluctuating speed/pitch of playback) creating the only polyphony in the recorded component.

In the *Closing* section the entire recording slows gradually to a complete stop, resulting in the generation of gradually lower and darker noteheads. The sections follow the same timing and order in each performance regardless of the audio content.

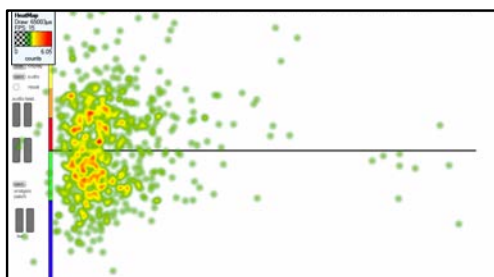
### 3. SEMANTIC PRINCIPLES

In any generative, real-time score, “semantic soundness” is an imperative for efficiently communicating notational aims to the performer. The conceit of the traditional beam, stem and notehead is employed as it is presumed to be already familiar to the performer. However, the generative score also differs from a traditional score in a number of ways.

The score scrolls from right to left on a *jit.lcd* screen with the dimensions 1300x700px. Graphics are drawn on the right side of the screen at the horizontal 1200 px position. The screen is scrolled leftwards at a rate of 100 px/s. Previous research on the semantic issues of screenscores had suggested that this is a middle-range scroll rate [3, 8, 9] roughly equivalent to reading a traditional score at a moderate speed. At play here is the trade-off between excessive note density (at a slow scroll-rate) and the limits of the eye’s fixation rate (at a fast scroll-rate) [10, 11, 12, 13].

The performer executes events in the notation at the point they reach a line (play-head) on the left-hand side (0 px) of the screen. The graphics take 12s, from the point of drawing, to reach to left-edge of the screen and the audio that generated them is also delayed for 12s. Graphics reach the play-head simultaneously with the audio events that caused their generation. In this manner, the performer has a visualisation of what will occur in the recording 12s in advance of it being heard. The process gives the performer a significant amount of time to view and access the up-coming passages of music, significantly greater than the eye-hand span (the distance between the point of fixation and the point of performance) for music readers suggested by sight-reading studies (between 1.5 and 3 cm) [14].

Eye-tracking tests [15] were conducted upon a number of participants reading *SoR*. The results indicate that the majority of performer fixations do indeed fall within the first 3 cm of the score, however there were a significant number of examples of “look-ahead” fixations spanning the entire score (See Figure 5).



**Figure 5.** Eye-tracking heat-map of *Semantics of Redaction*. The colours indicate red as areas of most frequent fixation and green as least frequent fixation.

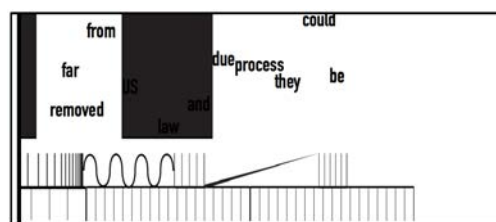
Although it is possible to draw other notehead shapes, (including more traditional looking rounded noteheads), the rectangle is ideal to indicating a precise onset of the performer’s note execution. As previously mentioned, to aid clarity in the correspondence between the spatial distribution of the note-events and their temporal distribution, note stems always occur on the left hand side of the notehead.

Frequency data is converted to just five HSL Hue values: Yellow (165), Orange (90), Red (0), Green (330), Blue (670). The ordering is suggested by research conducted at the University of California Berkeley, which shows that the brighter-darker hue continuum is strongly correlated with higher-lower frequency continuum [16, 17], due to what has been termed Weak Synaesthesia [18] or Crossmodal Correspondence [19]. Saturation and Luminance (and vertical position) are all continuously scaled with a potential resolution of 700 degrees (the number of vertical pixels).

The additional graphic symbols discussed in the previous section are intended to imply natural, heuristic correspondences both gesturally and sonically, due to “natural constraints on the ways in which sounds are mapped on to objects” [20]. In this way the graphic adjacent to *Commentary I* in Figure 4 should imply to the performer a wave-like gesture/sound increasing in magnitude.

Interestingly, because the brightness and noisiness of accents are represented visually, there are also some observable correspondences between phonemes and the saturation and luminance of noteheads: for example, sibilant sounds tend to be brighter and less saturated in the current system. This may prove to be an interesting phenomena to explore in future work.

The first versions also inverted the function of speech: text was printed onto to the score and intended to be either read by the performer or by the computer using Masayuki Akamatsu’s *aka.speech* object (See Figure 6).



**Figure 6.** An early version of *The Semantics of Redaction*.

The final version of the notation - where vertical space approximates and notehead colour approximately to frequency and to instrument family represents the natural fluid, non-metrical, rhythms of speech very well. Small fluctuations in vertical height were readily interpreted as the huge range of strikes and strokes and scrapes that can be coaxed from percussion instruments and other objects.

The performance practice of muting the speech and allowing an instrumental interlude is similar to the approach taken by Alvin Lucier in his work *Carbon Copies* (1989), in which a field recording is used as an “audio-score” for performers that is heard alone, then in addition to the performer(s) who are emulating it, is muted so the audience only hear the performer and then muted even for the performer leaving them to improvise an emulated version of the recording.

The performance is perhaps best situated at the nexus between interpretation and improvisation, in which the performer ranges between precise interpretation of symbols and more metaphorical improvisation in response to graphical notation.

#### 4. CONCLUSIONS

*The Semantics of Redaction* is a flexible vehicle for a performer to explore audio from the media-world in a near-real-time manner. The author has also performed the work on a number of occasions on bass clarinet, reading the colour-coded notation as five instrumental register rather than instrumental families. Multiple performers have also successfully performed the work. It is possible to perform with multiple networked scoreplayers, allowing for interaction with visualisations that focus on varied frequency, amplitude and timbral parameters of the same recording, although this has not yet been attempted.

#### Acknowledgments

Research on this work and into notational semantics, eye-tracking and screenscores was funded by an Early Career Researcher Grant from Edith Cowan University. The Semantics of Redaction was commissioned by Vanessa Tomlinson with funding from the Australia Council for the Arts. Early experiments were conducted at the *Orpheus Instituut for Advanced Studies & Research In Music*. Very special thanks to percussionist Vanessa Tomlinson for the dialogue that led to the creation of this work.

#### 5. REFERENCES

- [1] Tomlinson, V. (2014). 8Hits. <http://8hits.org/>
- [2] Klein, N. (2014). *This Changes Everything: Capitalism versus the Climate*, New York: Simon & Schuster.
- [3] Vickery, L. (2014). Exploring a Visual/Sonic Representational Continuum. Proceedings of the International Computer Music Conference, Athens, Greece. pp. 177-184.
- [4] Vickery, L. (2004). Interactive control of higher order musical structures”, in Proceedings of the Australasian Computer Music Conference 2004, Victoria University, Wellington New Zealand.
- [5] Vickery, L. (2003). Non-linear structures for real-time interactive musical works. Australasian Computer Music Conference. Vickery, L.. Perth, WA. Australasian Computer Music Association. 2003: 87-94
- [6] Jehan, T. (2004). Event-synchronous music analysis/synthesis. In Proc. Digital Audio Effects Workshop (DAFx), Naples, Italy.
- [7] Vickery, L. (2012). Adapting John Cage’s Radio Music for a digital score player. Proceedings of the *Australasian Computer Music Conference 2012* 69-77. Griffith University, Brisbane, Australia.
- [8] Vickery, L. (2014). The Limitations of Representing Sound and Notation on Screen. *Organised Sound* 19(3). Cambridge University Press.
- [9] Vickery, L. (2014). Notational Semantics in Music Visualisation and Notation. Proceedings of the *Australasian Computer Music Conference 2014* Melbourne, Victoria.
- [10] T. Goolsby, “Eye-movement in music reading - Effects of reading ability, notational complexity, and encounters”. *Music Perception*, vol.12, pp. 77-96, 1994.
- [11] F. Truitt, C. Clifton, A. Pollatsek, & K. Rayner, “The perceptual span and eye-hand span in sight reading music”. *Visual Cognition*, vol. 4, pp. 143– 161, 1997.
- [12] A. Waters, E. Townsend, & G. Underwood, “Expertise in musical sight-reading: A study of pianists”. *British Journal of Psychology*, vol. 89, pp. 123-149, 1998.
- [13] A. Waters, G. Underwood, & J. Findlay, “Studying expertise in music reading: Use of a pattern- matching paradigm”, *Perception and Psychophysics*, vol. 59, no. 4, pp. 477-488, 1997.
- [14] Gilman E. & Underwood, G. (2003). Restricting the field of view to investigate the perceptual spans of pianists”. *Visual Cognition*, vol. 10, no. 2, pp. 201-232, 2003.
- [15] The experiments were conducted using the Tobii X2-30 eye-tracker.
- [16] Griscom, W. S., and Palmer, S. E. (2013). Cross-modal Sound-to-Sight Associations with Musical Timbre in Non-Synesthetes. *13th Annual Meeting of the Vision Science Society*, May 10 -15, Naples, FL.
- [17] Griscom, W. S., and Palmer, S. E. (2012). The Color of Musical Sounds in Non-Synesthetes. *12th Annual Meeting of the Vision Science Society*, Naples, FL.
- [18] Martino, G., and Marks, L. E. (2001). Synesthesia: Strong and weak. *Current Directions in Psychological Science*, 10, 61–65.
- [19] Deroy O, Spence C. (2013) Weakening the case for ‘weak synaesthesia’: Why crossmodal correspondences are not synaesthetic. *Psychonomic Bulletin & Review* 20: 643–664.
- [20] Ramachandran, V.S., and Hubbard, E.M. (2001). Synaesthesia - A Window Into Perception, Thought and Language. *Journal of Consciousness Studies*, 8, No. 12, 3–34: 19.

## An Approach to the Generation of Real-Time Notation via Audio Analysis: *The Semantics of Redaction*

Lindsay Vickery

Edith Cowan University

l.vickery@ecu.edu.au

### ABSTRACT

*This paper discusses approach mechanically and conceptually to generation notation in real-time in MaxMSP in the work The Semantics of Redaction. The work uses analysis data to generate scrolling, but conventional looking notation. The issues discussed include scaling and mapping of audio data to visual representation, the formal structure of the work and its relation to the generative processes, semantic principles guiding the notational concept and the development of a performance practice for the work. The intention of the work is to provide a flexible vehicle for a performer to explore audio from the media-world in a near-real-time manner using a recent, topical speech recording as the generative artifact for each performance.*

### 1. INTRODUCTION

*The Semantics of Redaction* [2014] (SoR) is a work for solo performer, pre-recorded audio and generative score. To perform the work, a recent, topical speech recording is chosen and loaded. The intent here is that the performer chooses a new recording for each performance emphasizing the “near real-time”, “daily news”-like aspect of the process. The audio is played through the work’s analysis patch and the analysis parameters (frequency, amplitude, brightness, noisiness and attack) are scaled to provide desirable outputs, in terms of range, density and diversity of the generated notation. Instrumentation is chosen by the performer as a commentary upon the subject matter of the recording. It is intended that the generative score is simple enough to perform after a brief period of familiarisation.

The work was written for percussionist Vanessa Tomlinson and officially premiered in her program 8 Hits [1], an ABC Classic FM live broadcast on November 1 2014. Tomlinson chose a speech made by climate-change activist Naomi Klein as part of the promotional tour for her book *This Changes Everything* [2014] [2].

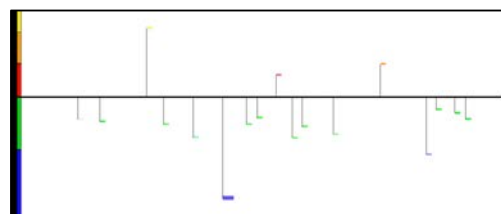
Copyright: © 2015 Lindsay Vickery. This is an open-access article distributed under the terms of the [Creative Commons Attribution License 3.0 Unported](https://creativecommons.org/licenses/by/3.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Along with its sister work *Lyrebird* [3], *SoR* began development during Tomlinson’s residency at the *Orpheus Institute for Advanced Studies & Research In Music* in December 2013.

Tomlinson is both a strong improviser and reader and the work specifically aimed to open a space in which she could interact with a chosen recording with a combination of precision and freedom. Like previous “series” by this composer, *Delicious Ironies* [2001] [4], *Splice* [2002] [5] and *Lyrebird* [2014] *SoR* is intended to be a structured improvisation environment that takes on different qualities in each performance.

### 2. GENERATIVE APPROACH

The notation for *SoR* (see Figure 1) is principally created by using accents detected in a speech recording in real-time to generate graphical symbols of varying vertical position, size and colour, determined by the frequency, amplitude and timbre of the recording at the accent point.



**Figure 1.** The opening section of *The Semantics of Redaction* [2014].

Analysis of the recording is derived from Tristan Jehan’s [6] *analyzer~* object, using attack, frequency, amplitude, brightness and noisiness data. The flow of analysis data is captured at the point of each detected attack. By default, the values defining an attack are an increase in the incoming signal of +10db in a period less than 100ms, however the performer may alter these values to achieve a better score/visualisation result.

The data is mapped in the following manner:

Frequency: Vertical height and notehead colour hue  
Amplitude: Notehead size  
Brightness: Notehead colour saturation  
Noisiness: Notehead colour luminance

The data is drawn to MaxMSP's *jit.lcd* object using the standard *paintrect* command. A stem is also drawn on the left side of the notehead above a central "beam". A left side stem is used regardless of whether the notehead is above or below the beam as an aid to visualising the precise onset of the accent and therefore performer's note execution.

Because of the variation of the acoustic characteristics of potential recordings and performer preferences, an analysis window allows the performer to scale incoming data (See Figure 2). Scaling changes may be necessary because of variations in the average frequency of the speaking voice or ambient noise in the recording. From the performer's perspective the attack rate and sensitivity values create changes in the density of note events in the score.

The analysis window provides the performer with "peak", "trough" and "mean" values for each of the analysis parameters from the recording, as well as a graphical display showing the most recent 100 values for each parameter. A switch graphically displays the variations in notehead colours. A *preset* object allows the performer to save the chosen scalings for each new recording.

Although it would be possible to automate the scaling process, allowing the performer to set the values provides a means to individualise the score that is generated by each recording: for example to achieve variations in the number of noteheads that fall within different "registers".

Unlike *Lyrebird*, which has a range of colour interpolations and generates a spectrogram-like score, only five hues (Yellow, Orange, Red, Green and Blue) are used in *SoR* (although there is continuous variation in the saturation and luminance of each colour). This approach was taken to achieve a clear differentiation between five registers in the score, which are conceived as instrument groups by the percussionist: for example ranging from yellow (high-pitched instruments) to blue (low-pitched instruments). The notehead size is intended to correlate to dynamic range.

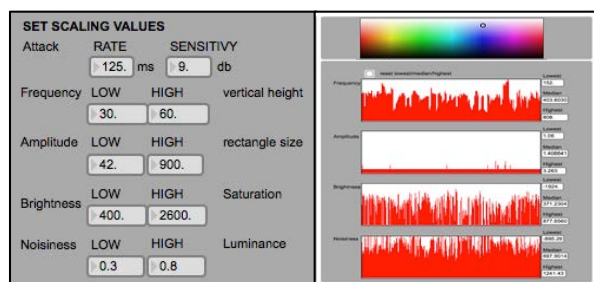


Figure 2. The Analysis window.

In addition to the generative process described above - which occur during the *Opening* and *Body* sections of the work - there are three other modes: *Commentary*, *Interlude* and *Redactio*. These modal changes form sections in the

structure of the work. In each of these sections graphical symbols are added to the notation. Figure 3 gives an indication of the morphologies of the "families" of symbols. There are 76 symbols in total.



Figure 3. Families of Graphical Symbols employed in different sections of *The Semantics of Redaction*.

In the *Commentary* sections the noteheads are extended indicating the note event should be held proportionally for that duration. The wave-like black lines that widen or narrow, indicate timbral or dynamic variation of the note that remains constant, (See Figure 4a.)

In the *Interlude* section some noteheads appear without stems indicating non-accented/slurred gracenotes, rather like liquescent neumes in medieval notation, while stems indicate accents. Graphical symbols in the form of black vertical lines are added and are interpreted as note or noise repetitions (See Figure 4b.). During this section the performer plays solo: the recorded speech is not heard by the audience (although it still drives the generation of notation). Audio from the speech is recorded into buffers that are heard during the "redactio" sections.

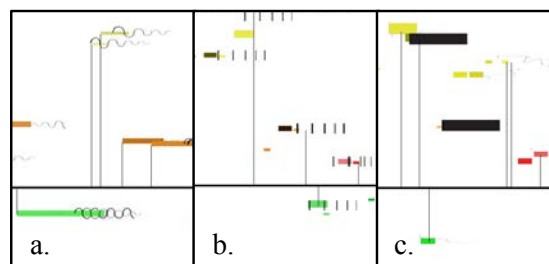


Figure 4. *Commentary* (4a), *Interlude* (4b) and *Redactio* (4c) sections from *The Semantics of Redaction*.

The *Redactio* section (See Figure 4c.) combines elements of the *Body* and *Commentary* sections, however black rectangles (redactangles) accompanied by beating sine tones, obscure elements of the score (and recording). Prerecorded samples from the previous *Interlude* are manipulated (via fluctuating speed/pitch of playback) creating the only polyphony in the recorded component.

In the *Closing* section the entire recording slows gradually to a complete stop, resulting in the generation of gradually lower and darker noteheads. The sections follow the same timing and order in each performance regardless of the audio content.



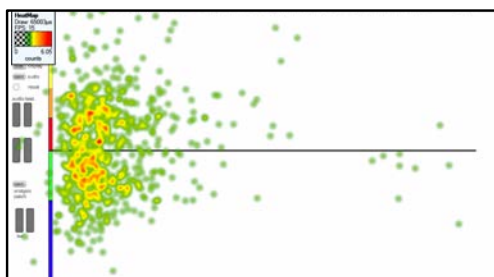
### 3. SEMANTIC PRINCIPLES

In any generative, real-time score, “semantic soundness” is an imperative for efficiently communicating notational aims to the performer. The conceit of the traditional beam, stem and notehead is employed as it is presumed to be already familiar to the performer. However, the generative score also differs from a traditional score in a number of ways.

The score scrolls from right to left on a *jit.lcd* screen with the dimensions 1300x700px. Graphics are drawn on the right side of the screen at the horizontal 1200 px position. The screen is scrolled leftwards at a rate of 100 px/s. Previous research on the semantic issues of screenscores had suggested that this is a middle-range scroll rate [3, 8, 9] roughly equivalent to reading a traditional score at a moderate speed. At play here is the trade-off between excessive note density (at a slow scroll-rate) and the limits of the eye’s fixation rate (at a fast scroll-rate) [10, 11, 12, 13].

The performer executes events in the notation at the point they reach a line (play-head) on the left-hand side (0 px) of the screen. The graphics take 12s, from the point of drawing, to reach to left-edge of the screen and the audio that generated them is also delayed for 12s. Graphics reach the play-head simultaneously with the audio events that caused their generation. In this manner, the performer has a visualisation of what will occur in the recording 12s in advance of it being heard. The process gives the performer a significant amount of time to view and access the up-coming passages of music, significantly greater than the eye-hand span (the distance between the point of fixation and the point of performance) for music readers suggested by sight-reading studies (between 1.5 and 3 cm) [14].

Eye-tracking tests [15] were conducted upon a number of participants reading *SoR*. The results indicate that the majority of performer fixations do indeed fall within the first 3 cm of the score, however there were a significant number of examples of “look-ahead” fixations spanning the entire score (See Figure 5).



**Figure 5.** Eye-tracking heat-map of *Semantics of Redaction*. The colours indicate red as areas of most frequent fixation and green as least frequent fixation.

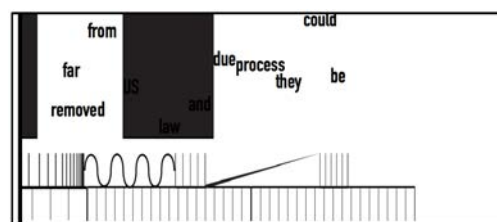
Although it is possible to draw other notehead shapes, (including more traditional looking rounded noteheads), the rectangle is ideal to indicating a precise onset of the performer’s note execution. As previously mentioned, to aid clarity in the correspondence between the spatial distribution of the note-events and their temporal distribution, note stems always occur on the left hand side of the notehead.

Frequency data is converted to just five HSL Hue values: Yellow (165), Orange (90), Red (0), Green (330), Blue (670). The ordering is suggested by research conducted at the University of California Berkeley, which shows that the brighter-darker hue continuum is strongly correlated with higher-lower frequency continuum [16, 17], due to what has been termed Weak Synaesthesia [18] or Crossmodal Correspondence [19]. Saturation and Luminance (and vertical position) are all continuously scaled with a potential resolution of 700 degrees (the number of vertical pixels).

The additional graphic symbols discussed in the previous section are intended to imply natural, heuristic correspondences both gesturally and sonically, due to “natural constraints on the ways in which sounds are mapped on to objects” [20]. In this way the graphic adjacent to *Commentary I* in Figure 4 should imply to the performer a wave-like gesture/sound increasing in magnitude.

Interestingly, because the brightness and noisiness of accents are represented visually, there are also some observable correspondences between phonemes and the saturation and luminance of noteheads: for example, sibilant sounds tend to be brighter and less saturated in the current system. This may prove to be an interesting phenomena to explore in future work.

The first versions also inverted the function of speech: text was printed onto to the score and intended to be either read by the performer or by the computer using Masayuki Akamatsu’s *aka.speech* object (See Figure 6).



**Figure 6.** An early version of *The Semantics of Redaction*.

The final version of the notation - where vertical space approximates and notehead colour approximately to frequency and to instrument family represents the natural fluid, non-metrical, rhythms of speech very well. Small fluctuations in vertical height were readily interpreted as the huge range of strikes and strokes and scrapes that can be coaxed from percussion instruments and other objects.

The performance practice of muting the speech and allowing an instrumental interlude is similar to the approach taken by Alvin Lucier in his work *Carbon Copies* (1989), in which a field recording is used as an “audio-score” for performers that is heard alone, then in addition to the performer(s) who are emulating it, is muted so the audience only hear the performer and then muted even for the performer leaving them to improvise an emulated version of the recording.

The performance is perhaps best situated at the nexus between interpretation and improvisation, in which the performer ranges between precise interpretation of symbols and more metaphorical improvisation in response to graphical notation.

#### 4. CONCLUSIONS

*The Semantics of Redaction* is a flexible vehicle for a performer to explore audio from the media-world in a near-real-time manner. The author has also performed the work on a number of occasions on bass clarinet, reading the colour-coded notation as five instrumental register rather than instrumental families. Multiple performers have also successfully performed the work. It is possible to perform with multiple networked scoreplayers, allowing for interaction with visualisations that focus of varied frequency, amplitude and timbral parameters of the same recording, although this has not yet been attempted.

#### Acknowledgments

Research on this work and into notational semantics, eye-tracking and screenscores was funded by an Early Career Researcher Grant from Edith Cowan University. The Semantics of Redaction was commissioned by Vanessa Tomlinson with funding from the Australia Council for the Arts. Early experiments were conducted at the *Orpheus Instituut for Advanced Studies & Research In Music*. Very special thanks to percussionist Vanessa Tomlinson for the dialogue that lead to the creation of this work.

#### 5. REFERENCES

- [1] Tomlinson, V. (2014). 8Hits. <http://8hits.org/>
- [2] Klein, N. (2014). *This Changes Everything: Capitalism versus the Climate*, New York: Simon & Schuster.
- [3] Vickery, L. (2014). Exploring a Visual/Sonic Representational Continuum. Proceedings of the International Computer Music Conference, Athens, Greece. pp. 177-184.
- [4] Vickery, L. (2004). Interactive control of higher order musical structures”, in Proceedings of the Australasian Computer Music Conference 2004, Victoria University, Wellington New Zealand.
- [5] Vickery, L. (2003). Non-linear structures for real-time interactive musical works. Australasian Computer Music Conference. Vickery, L.. Perth, WA. Australasian Computer Music Association. 2003: 87-94
- [6] Jehan, T. (2004). Event-synchronous music analysis/synthesis. In Proc. Digital Audio Effects Workshop (DAFx), Naples, Italy.
- [7] Vickery, L. (2012). Adapting John Cage’s Radio Music for a digital score player. Proceedings of the *Australasian Computer Music Conference 2012* 69-77. Griffith University, Brisbane, Australia.
- [8] Vickery, L. (2014). The Limitations of Representing Sound and Notation on Screen. *Organised Sound* 19(3). Cambridge University Press.
- [9] Vickery, L. (2014). Notational Semantics in Music Visualisation and Notation. Proceedings of the *Australasian Computer Music Conference 2014* Melbourne, Victoria.
- [10] T. Goolsby, “Eye-movement in music reading - Effects of reading ability, notational complexity, and encounters”. *Music Perception*, vol.12, pp. 77-96, 1994.
- [11] F. Truitt, C. Clifton, A. Pollatsek, & K. Rayner, “The perceptual span and eye-hand span in sight reading music”. *Visual Cognition*, vol. 4, pp. 143– 161, 1997.
- [12] A. Waters, E. Townsend, & G. Underwood, “Expertise in musical sight-reading: A study of pianists”. *British Journal of Psychology*, vol. 89, pp. 123-149, 1998.
- [13] A. Waters, G. Underwood, & J. Findlay, “Studying expertise in music reading: Use of a pattern- matching paradigm”, *Perception and Psychophysics*, vol. 59, no. 4, pp. 477-488, 1997.
- [14] Gilman E. & Underwood, G. (2003). Restricting the field of view to investigate the perceptual spans of pianists”. *Visual Cognition*, vol. 10, no. 2, pp. 201-232, 2003.
- [15] The experiments were conducted using the Tobii X2-30 eye-tracker.
- [16] Griscom, W. S., and Palmer, S. E. (2013). Cross-modal Sound-to-Sight Associations with Musical Timbre in Non-Synesthetes. *13th Annual Meeting of the Vision Science Society*, May 10 -15, Naples, FL.
- [17] Griscom, W. S., and Palmer, S. E. (2012). The Color of Musical Sounds in Non-Synesthetes. *12th Annual Meeting of the Vision Science Society*, Naples, FL.
- [18] Martino, G., and Marks, L. E. (2001). Synesthesia: Strong and weak. *Current Directions in Psychological Science*, 10, 61–65.
- [19] Deroy O, Spence C. (2013) Weakening the case for ‘weak synaesthesia’: Why crossmodal correspondences are not synaesthetic. *Psychonomic Bulletin & Review* 20: 643–664.
- [20] Ramachandran, V.S., and Hubbard, E.M. (2001). Synaesthesia - A Window Into Perception, Thought and Language. *Journal of Consciousness Studies*, 8, No. 12, 3–34: 19.