

A Language Apparatus

Presented at the Electronic Literature Organization Conference, Bergen, Norway, 2015

Abstract

The author discusses the collaborative creative projects Bodytext, Tower and Crosstalk, exploring how language and communication function in a context where human and machine are each involved, in a co-dependent manner, in the creation and interpretation of texts. The dynamics of a hybrid reading-writing apparatus are analysed to allow insights into the production and reception of meaning within a socio-technical system.

The three projects, Bodytext, Tower and Crosstalk, are language based digitally mediated performance installations. They each use progressive developments of generative and interpretative grammar systems.

Introduction

Computing, and its associated technologies, has changed language as profoundly as writing and printing before it. As a symbolic machine, a system of signs that reflexively operates upon and modifies itself, both carrying and making meaning, the principle of computation presents a shift in our linguistic and symbolic modalities, with subsequent implications for social and individual ontology. We have rapidly adopted the computer as personal companions, as extensions of ourselves. Many of us are soft-wired into the machine and the possibility of hard-wiring is being explored by artists and scientists. The computer, as a language system, has become part of us and we have become part of it.

The early Artificial Intelligence researcher Terry Winograd proposed that the computer is a new form of writing – not a medium for writing but writing itself, reformulated. Winograd's argument is premised on a theory of computation that understands computing to be a recursive symbolic process, where a symbolic system (language) acts upon itself to create new instances of the system - effectively, writing that writes itself. Within this framework the computer is conceived, in human experience, as a novel form of language – language that is auto-semiotic. Winograd observed that "the computer is a physical embodiment of the symbolic calculations envisaged by Hobbes and Leibniz. As such, it is really not a thinking machine, but a language machine" (Winograd 1991). Winograd's central argument is that computational processes are intrinsically symbolic and therefore language per se. His argument is that the computer is an evolution of writing, where that writing can be autonomic.

If we accept Winograd's argument then we can also accept that the computer can be considered to have changed language as profoundly as writing and printing before it. As a symbolic machine, a

system of signs that reflexively operates upon and modifies itself, both translating and generating meaning, the computer represents a new linguistic modality. However, again, we must remember that this is less about technical development and more to do with human revelation. What we see here is an emergent apprehension of the human, an evolving ontology for something that already exists. In this respect the computer is itself a representation - a model of how we understand what it is to be human. Intrinsic to this is an understanding of the human as premised on language, not dissimilar to the computer. What would it be like to be a human without language? Would we be human as we understand that concept - individuated social beings able to interact with other individuated social beings with the degrees of symbolic commonality and exchange (community and communication) such interaction requires.

The concept of 'social media' can, in this context, assume new meaning, suggesting that the computer is an intrinsic part of our contemporary cultural and social forms. If we exist and are primarily defined as social beings, enmeshed in community and culture, and if the computer (as an instance of language) has become central in the relations these concepts describe then the term 'social media' can be considered to denote not only a medium through which to socialise but a mediator (or, perhaps, more precisely, re-mediator) of the social. That is, in short, society is as much about computing as it is about anything else. This was a concept that informed the Remediating the Social conference, exhibition and accompanying book, where it was argued that "creativity might be considered a property emergent from a multi-modal social apparatus rather than, as is more commonly assumed, an attribute of individual or collective human agency" within the context of "media that are part of the apparatus we can identify as the social in action" (Biggs 2012).

It is salient, in this context, to consider N. Katherine Hayles suggestion that humans and technology are co-evolving, a process she terms technogenesis (Hayles 2012). Hayles recognises that the machinic element, particularly in the form of the internet, is now so sophisticated and explicitly symbolic, conditioning how we perceive the human-machine relationship as not only physical but also linguistic and social, that we can no longer deny our technologised condition. We are, in this sense, self-knowingly the machine. It is, in this light, an implicit understanding that the human-machine assemblage is primarily a social construct.

Whilst Hayles identifies this as a contemporary condition it can be argued that this is a process of co-evolution that has been underway for as long as humans have been able to communicate through the use of symbolic systems. If the computer is an evolution of language then it is reasonable to assume that what is the case in computing is also the case, more generally, for language. It is arguable that this process of progressive hybridisation has been underway for millennia and that it is, as such, an ancient defining characteristic of the human - that we have been socio-technical hybrids as long as we have been human. Andy Warhol's observation that

"everybody should be a machine", a less famous statement from the same interview where he famously stated "I want to be a machine" (Swenson 1963), can be considered less an aspirational statement than one of fact, particularly relevant as Warhol made this comment whilst discussing how community formation functions in different socio-political contexts (communism and capitalism) to drive social conformity.

Three artworks

The three projects, Bodytext, Tower and Crosstalk, are language based digitally mediated performance installations that reflect upon and explore, through performance, the ontologies of language-based beings who are components of a language-based socio-technical system. Each work employs progressive developments of generative and interpretative grammar systems. Bodytext (2010) was authored in Adobe Director and coded in Lingo and C++. Tower (2011) was developed for a large scale immersive virtual reality system and coded in Python. Crosstalk (2014) was developed and coded in Processing.

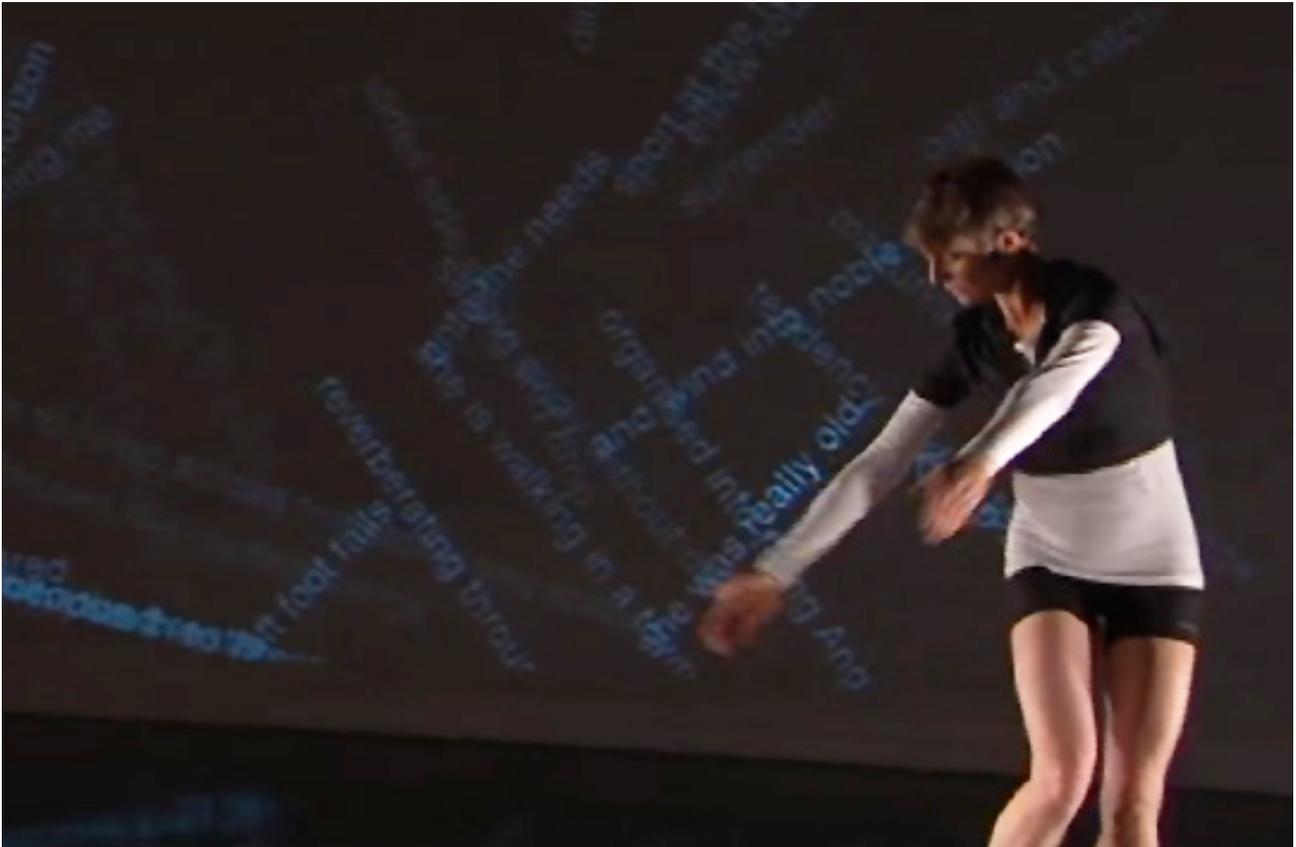
All these works engage the performer and audience as readers and writers enmeshed within reading and writing systems that function through a dispersed agency, much of which is not human but machine-based. Here there are no actors or actants (to paraphrase Bruno Latour) but only interactors, whether they be human or machine - they are all manifest as elements in a symbolic system, mutually dependent on one another.

It is worth noting here that the term actant does not have its origins in Actor-Network-Theory (Latour 2005) but in the semiotic theory of Algirdis Julien Greimas and, specifically, in his analysis of story-telling and the roles of characters within them - what Greimas called the actantial model (Greimas 1966). Within this framework fictional agents are defined as antagonists and protagonists, 'good-guys' and 'bad-guys', with the dynamics of the narrative configured and played-out through the relationships between the agents. In this respect, whilst the Actor-Network-Theory approach can be seen as relevant to the works discussed here it is possible that Greimas's work is the more effective for offering insights into how the human and non-human interactors in these artworks function as antagonists and protagonists, determining the poetic dynamics of the performances and interactions that comprise the works.

Bodytext

Bodytext (Biggs, Hawksley & Paine 2010) is an interactive performance work involving speech, movement and the body. A dancer's movement and speech are re-mediated within an augmented environment employing real-time motion tracking, voice recognition, interpretative language systems, projection and audio synthesis. The acquired speech, a description of an imagined dance, is re-written through projected digital display and sound synthesis, the performer causing

texts to interact and recombine with one another. What is written is affected by the dance whilst the emergent texts determine what is danced. The work questions and seeks insight into the relations between kinesthetic experience, memory, agency and language.



<http://littlepig.org.uk/installations/bodytext/index.htm>

In Bodytext the actantial dynamics that Greimas explores in fictional narrative are clearly evident, but here the interactors that drive this drama are neither fictional or character driven. Rather, the human and non-human interactors in Bodytext are the complement of one another, emergent within the dynamics of the system.

Visually the work is composed exclusively of texts, as spoken by the performer. The computer has a software model of all the words that might be used, in the form of a relatively indexed database of words (somewhat like a thesaurus) and can understand how the words connect to one another and, importantly, how they might alternately connect with one another. All the words uttered are data, in this sense, but also, due to their structural relations, possess syntactic meaning which the computer can understand. Some of the words are also commands, specific words linked to specific actions the software can enact, whether changing a conditional state in the system or triggering a particular event.

Each utterance by the performer becomes a software object itself, in what is an object-oriented software context. Each instance, or fragment, of text exists as an entity within the software and

also graphically, projected onto the video screen that immerses the performer. Just as the words the performer utters are gathered by the system and employed to construct objects so to are the performer's movements, which are then interact with and subsequently effect the physical behaviour of the graphical text objects. The outcome is that texts bounce off the performer's body, when in proximity to it, and subsequently interact with other text objects. When text objects pass through one another each text 'reads' the other text, employing an interpretative grammar system. Where the two texts find that they are able to switch syntactic elements without changing their syntactic form they do so, each text effectively re-writing the other (a generative grammar system). This results in texts appropriating words from one another and, subsequently, the emergence of new texts, in a form of recombination.

The performer, like the machine, also reads the texts, taking direction for the performance from them. Whilst the performer may have uttered the initial texts the texts that they then read and respond to, through their actions, can be quite distinct, leading to an evolving performance where what emerges is unknown, a consequence of the complex interactions within the system. The performer is required to not only interpret the evolving texts in their movement, with subsequent emergent interactions further evolving the texts, but to also read the texts out loud, with the effect that they are re-entered into and re-interpreted by the system. The increasing noise to signal ratio in the sonic dimension of the environment, detailed below, interferes with this process of human and machine 'reading' and 'writing', leading to novel elements entering into the system. Just as the sound system spins towards chaos and ultimate decay so does the linguistic content and structure of the space, causing the visual dimension of the work to become dense, convoluted and chaotic. The ultimate result is a nonsensical language-space that the performer can only interpret in a chaotic and demented manner, leading to their exhaustion and the physical, visual, audible and linguistic collapse of the system.

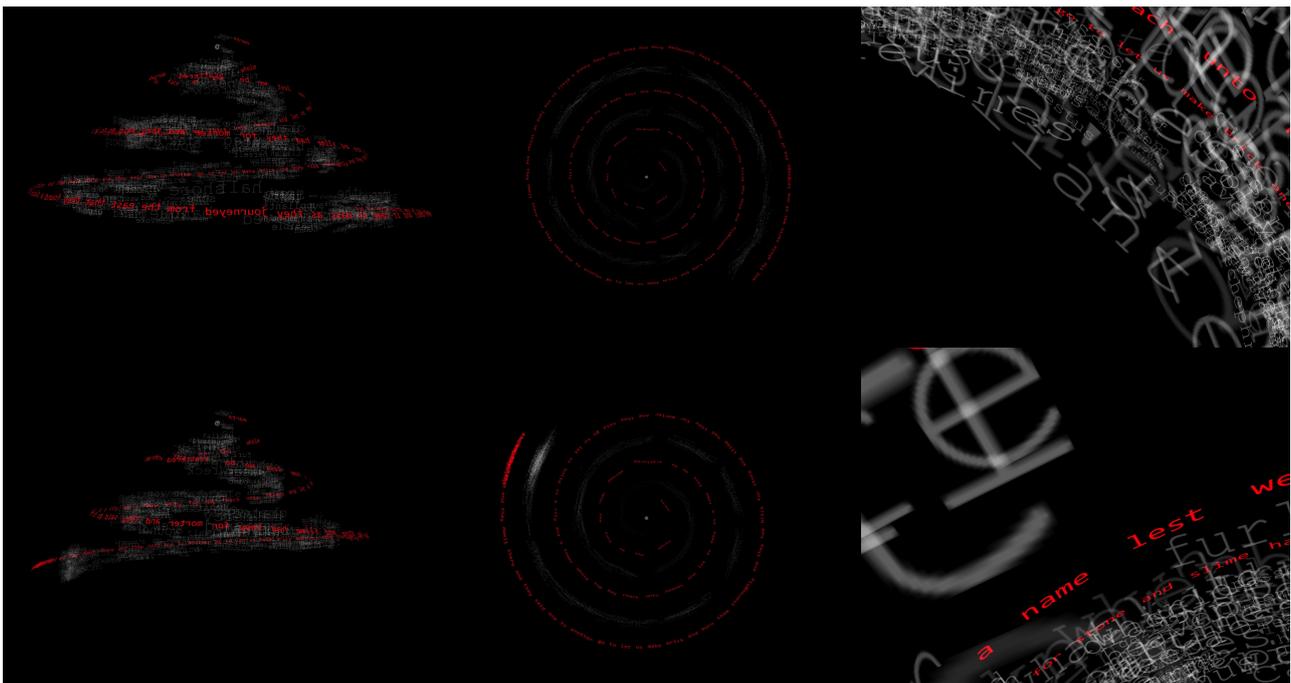
Whilst the textual fragments that initially 'seed' the generative and interpretative language system that underpins the software logic of the Bodytext interactive environment are based on the real-life events and experiences of the performer (dancer Sue Hawksley), and are delivered into the system through the speech of the performer, these texts are not defining of what unfolds in the work. The content of the texts is another agent in the determination of events but, as with all sources of agency in this work, the diversity of agency, the complexity of how they interact and the inherent 'noise' in the system all conspire to ensure that what ultimately manifests cannot be traced or apprehended relative to specific actions, symbolic interventions or fragments of data.

The noise in Bodytext is literally physically audible noise, which is pivotal, as the key antagonist, in defining the entropic dynamic of the work. The sound for the work, by composer Garth Paine, is acquired from the live speech and sounds of the physical movement of the performer and processed as a function of the performers actions, whether speech acts (specific words can effect

what sound processes are active) or movement (the relationship between parts of the performer's body set parameters for the sonic processing). The sound generated from and by the performer's physical movement is also picked up and resampled back into the sound system, creating a feedback loop that induces noise in the system. As the work unfolds, over a 30 minute duration, this noise increases (although not linearly - the profile of the noise ratio is dependent on the haptic interactions of performer and software) such that towards the end of the performance the noise within the sound system is greater than the signal. At this point the work breaks down and enters an entropic spiral that results in final stasis and the loss of all data in the system. The work ends in silence and darkness.

Tower

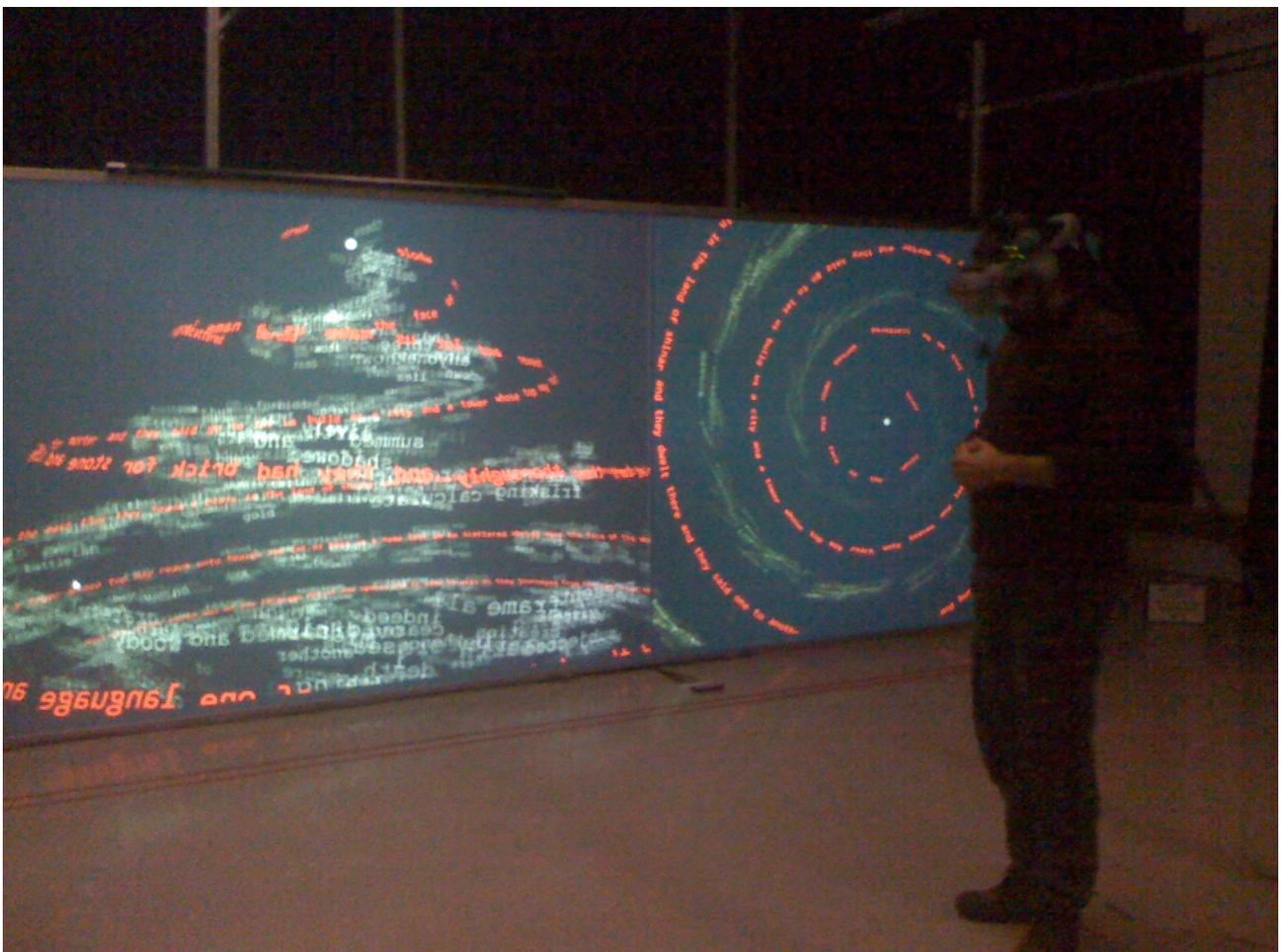
Tower (Biggs & Shovman 2011) is an interactive environment where the computer listens to and anticipates what is said by those interacting with it. A self-learning system, as the interactor speaks the computer displays what they say and the potential words they might speak next. The speaker may or may not use a displayed predictive word. New word conjunctions are added to the corpus employed for the statistical prediction process. Words uttered by the interactor appear as a red spiral of text, at the top of which the inter-actor is located within the virtual reality environment. Wearing a head-mounted-display the interactor can look wherever they wish, although they cannot move. The predicted words appear as white flickering clouds of text in and around the spoken words. What emerges is an archeology of speech where what is spoken can be seen amongst what might have been said, challenging the unique speaker's voice and our expectation and memory of events.



Tower is designed to be viewed in three different modes, each premised on a different point of view; 'first person', 'second person' and 'third person' - an approach borrowed from the conventions around the narrative voice in fiction. Interactors are free to choose which position they wish to view the work from, as all three are simultaneously available - although only one interactor can wear the HMD at any one time. Interactors can change their mode of viewing simply by moving to a different location or by passing the HMD on to another interactor.

The view experienced when wearing the HMD places the interactor at the centre of events, looking out at the text streams that surround them. Everything that comes into the virtual world comes from the mouth of this interactor. They see the words they speak emerging from out of their mouth, as they speak them, forming into a spiral below them which appears to support them in the vertiginous space. This is the view of the 'first person'.

Other interactors can stand around the wearer of the HMD and watch them interacting with and writing the work into being, immersed in the 3D projection of the virtual world. These interactors wear stereoscopic 3D glasses and see the projected word-spiral 'tower' in 3D. The scale of the projection and their near-field location relative to the screen places them in an immersive context, although one where they are not at the centre of the action. These interactors are in the position of the 'second person'.



Another element of the work is a two dimensional projection of the virtual world, but from above (or below - in this context these two positions are equivalent). In this projection the array of words is rendered as a flat graphic spiral, allowing all the words to appear as the same size. It is only in this view that it is possible to read all the words as a linear text. This is the 'third person' view.

Just as the work offers distinct modes of engagement so it also allows multiple textual readings. The interactor might choose to focus on the red spiral of text, composed of what has been said, or they could choose to read any number of the sub-spiral tributaries of semi-transparent white texts that flicker around the ever expanding red spiral of words. The textual context of these texts is formed through the statistical prediction of subsequent words as each word uttered by the 'first person' interactor appears at the top (or centre, depending on your point of view) of the spiral. The relative position, size and transparency of each predicted word is a function of its statistical likelihood of it appearing subsequent to the previous spoken word in the reference textual corpus. The reference corpus can be seeded (e.g.: taken from existing texts) or emerge as the outcome of all the words that have been spoken in the system - or both.

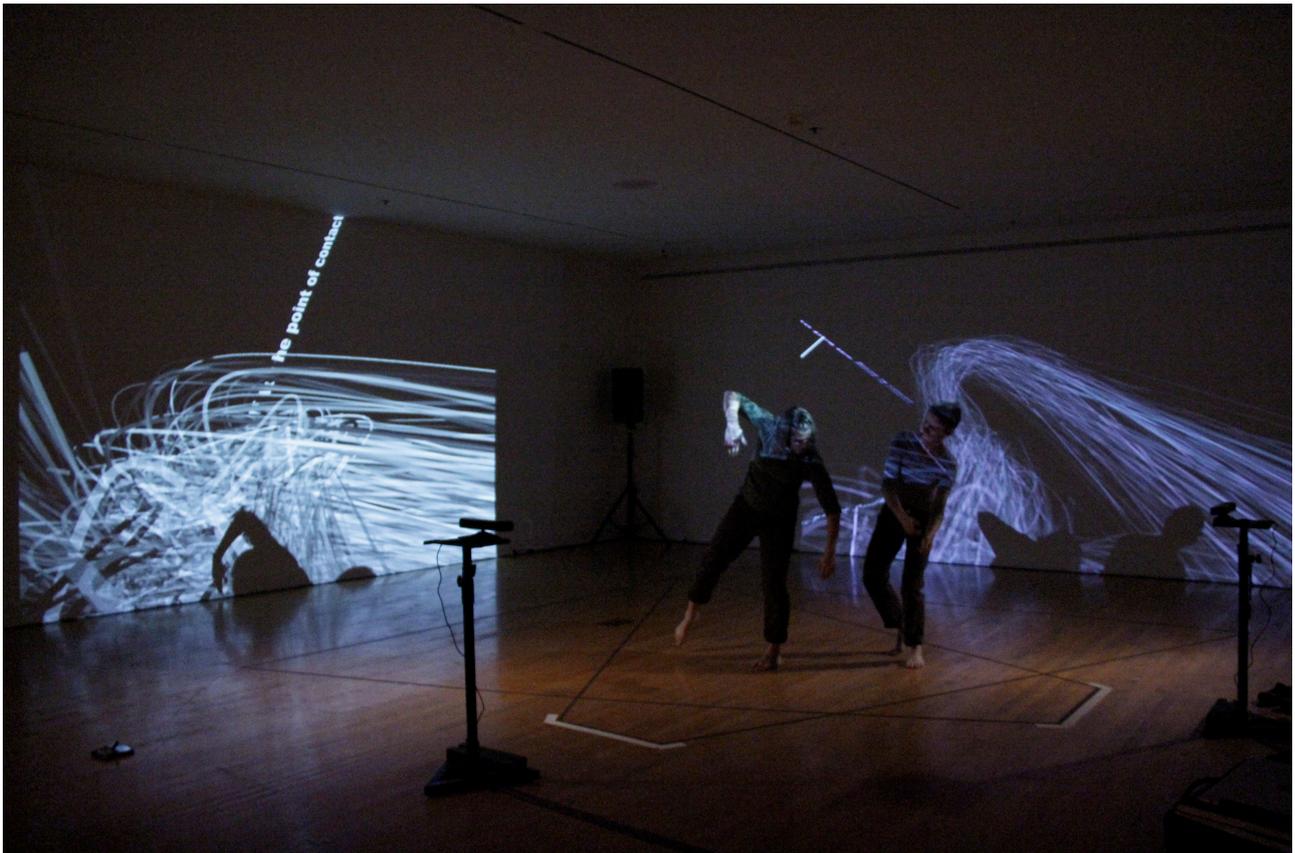
The sub-spirals of predicted text come to resemble flickering clouds, almost a fog, of text as they increase in number and density. The flickering quality of the texts is a function of the head movements of the 'first person' interactor. The HMD the 'first person' interactor wears is fitted with a motion sensor and this generates data that is employed to independently affect the rate, angle and degree of a multi-axis fluttering movement of each individual white text element (word). If the 'first person' interactor moves their head rapidly, perhaps turning around to view the words that were behind them, the white clouds of words that envelope the main red spiral of words become highly agitated, further emphasising their fog-like qualities.

The height and circumference of the spiral structure is the outcome of the number of words that have been spoken since the system was initialised and thus constantly grows in scale, the longer it is interacted with. For a 'third person' interactor this can lead to the number of the words in the spiral being so great that in order for the full scale of the spiral to be rendered in the projection the point-size of the text elements is so small as to render the text illegible. Indeed, after extended interaction the scale of the text is so small as to resemble a fine red spiral line enveloped in a soft flickering white mist. In this scenario the 'first person' interactor finds themselves at the top of a huge spiral structure that appears to disappear into what appears to be the nearly infinite depths of the swirling mist of words below them, suggesting an all-enveloping apparatus.

Crosstalk

'Crosstalk' (Biggs, Hawksley & Paine 2014a) is a multi-performer interactive environment where movement and speech are remediated within an augmented 3D environment employing real-time motion tracking, multi-source voice recognition, interpretative language systems, a bespoke

physics engine, large scale projection and surround-sound audio synthesis. Acquired speech of interactors is remediated through projected digital display and sound synthesis, the interactors physical actions causing texts to interact and recombine with one another. The elements in the system all affect how each adapts, from state to state as people, machines, language, image, movement and sound interact with one another. 'Crosstalk' explores social relations, as articulated in performative language acts, in relation to generative ontologies of self-hood and the capacity of a socio-technical space to 'make people'.

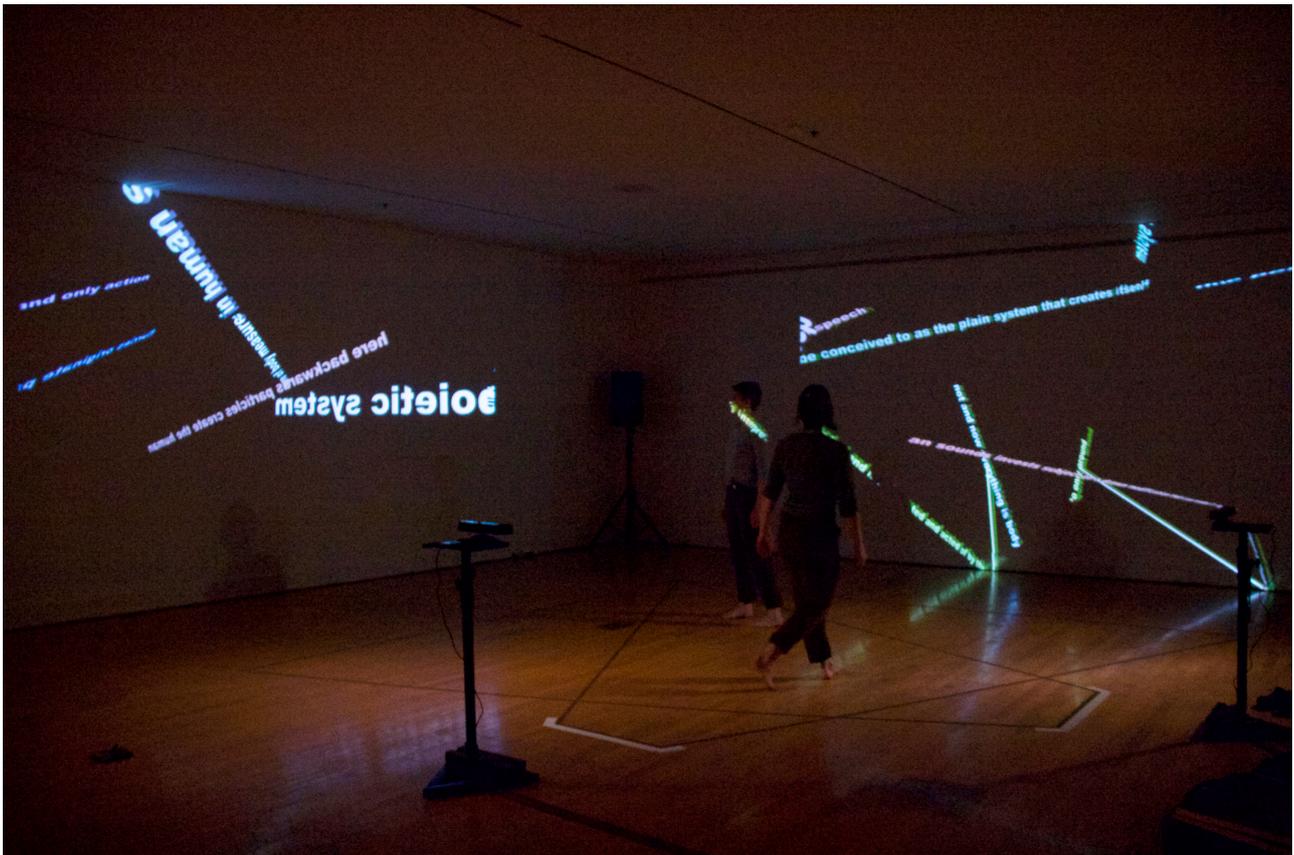


<http://littlepig.org.uk/installations/crosstalk/index.htm>

As outlined elsewhere, "Crosstalk developed out of the processes undertaken in Bodytext, specifically the 'drama of the performance', which arose from an antagonistic but interdependent human/machine relationship. Created for two performers, Crosstalk engages social relations as articulated through performative language acts. The project explores ontologies of self-hood within the generative potential of a technologically mediated social space" (Biggs, Hawksley & Paine 2014b). Similar to Bodytext, text is acquired by the system from the speech of the interactors, in this case two performers. Again, as in Bodytext, the interactors can physically interact with the texts, causing them to ricochet around the space and to collide and re-write one another, using an evolution of the interpretative and generative grammar engines employed in the earlier work. However, where Bodytext was rendered as a two dimensional projection, with all the text objects appearing on the same spatial plane, in Crosstalk the texts are rendered into a fully 3D

environment that is projected on the walls around the interactors, immersing them fully within the three dimensional space. Crosstalk also employs a fully featured physics engine, the parameters of which the interactors can control with their speech and movements (e.g.: turning gravity on or off or causing the text objects to be either repelled by or attracted to the interactors bodies).

Whereas in Bodytext the interactor was modeled as a two dimensional object within the system, with which the text objects could then interact, in Crosstalk the interactors bodies are modeled as fully three dimensional objects in a three dimensional interactive volume that is congruent with the actual physical performance space of the installation. In both works these interactor models are not visible in the projections.



As with Tower, and following on from the three dimensional quality of the work, the point of view in Crosstalk is multiple and dynamic. For much of the performance the point of view in the work is a default one; that of an imaginary 'third person' who can look into the three dimensional world of the work from an 'idealised' location outside the environment. Each of the projections that comprise the work is rendered from its own idealised 'third person' point of view and the interactors and the audience see the same thing, although from their respective locations relative to the projections. However, just as the interactors can control the parameters of the physics engine and the rules that govern the behaviour of the objects in the space, so too they can control the point of view from which the projections are rendered. With a single word they can cause the 'third person' point of view to be transformed to a 'first person' point of view, with the 'camera' (the position and

orientation from which the projections are rendered around the interactors) coming to be located on the head of the interactor who issues the command, oriented towards the head of the other interactor. Thus the point of view is that from the first interactors location, looking towards the second interactor. What is rendered in the corresponding projection for the first interactor is thus the outcome of the movement of the heads of the two interactors. Each interactor can independently control the camera corresponding to the projection associated with them, so that one or another of the projections, or both, can be rendered from a 'first' or 'third person' point of view. As the interactors are moving so to are the points of view, the three dimensional views around the interactors oscillating and shifting along an axis that is drawn between the heads of the two interactors.

In Crosstalk the visual world is not only composed of the text objects acquired from the speech of the interactors. The interactors can also control the visualisation of the virtual 3D models the system creates of the interactors to facilitate the interactions within the system. The interactors can make these models appear or disappear, but they can only do this when both interactors are in the interaction space. This is because what is rendered in the projections is not the model of a single interactor but a number of different models where each is based upon elements acquired from each of the interactors. The effect here is of emergent graphical agents comprised of distinct elements from each of the interactors, where the actions of the two interactors inform the action of each of the emergent models. This allows the interaction between the interactor-models and the text objects to be explicitly rendered, although the visual complexity tends to be such that causality remains glimpsed rather than clearly traced. This is especially the case when the interactors have turned on interactor-model rendering, heavily populated the environment with text objects and situated the 'cameras' on their heads. In these phases of the performance the visual character of the work is intense, composed of dense networks of oscillating texts and the dynamic visual flux of the interactor models, all rendered, and thus seen by the audience, from a fully immersive 'first person' point of view. This immersive experience is further enhanced by the surround sound, with the multi-point speaker system located around the perimeter of both the performance and audience space.

Conclusion

Three artworks, Bodytext, Tower and Crosstalk, have been discussed as examples of interactive environments that operate as language apparatus - systems where the actantial relations between human interactors, software systems, live speech, auto-semiotic texts and an evolving sound stage determine the evolution and outcome of the work, the interactors and other elements. In the dynamic and formative relations that emerge within each work we see the evolution of the system, with the subsequent problematisation of key attributes of individual agency, such as that of 'point of view', the origin of the performative and the hermeneutic. Each of these works is concerned with

self-knowledge and engages individual and collective ontologies, asking how we might imagine multiple states of individuation - states of being where we are fragmented, diffused and braided as manifold aspects of agency.

Simon Biggs, Adelaide, July 2015

Bibliography:

Biggs, S. (2012) *Remediating the Social*, Bergen: ELMCIP - University of Bergen.

Biggs, S, Hawksley, S & Paine, G. (2010) *Bodytext*,
<http://littlepig.org.uk/installations/bodytext/index.htm> (accessed 22.07.2015)

Biggs, S, Hawksley, S & Paine, G. (2014a) *Crosstalk*,
<http://littlepig.org.uk/installations/crosstalk/index.htm> (accessed 22.07.2015)

Biggs, S, Hawksley, S & Paine, G. (2014b) *Crosstalk: Making People in Interactive Spaces*,
Proceedings of the 2014 International Workshop on Movement and Computing, IRCAM, Paris,
New York: Association of Computing Machinery - D.O.I.: 10.1145/2617995.2618006

Biggs, S, & Shovman, M. (2011) *Tower*, <http://littlepig.org.uk/tower/index.htm> (accessed 22.07.2015)

Greimas, A. J. (1966) *Structural Semantics: An Attempt at a Method*, Lincoln (NE): University of Nebraska Press.

Hayles, N. K. (2012) *How We Think: Digital media and Contemporary Technogenesis*, Chicago: University of Chicago Press.

Latour, B. (2005) *Reassembling the Social: An Introduction to Actor-Network-Theory*, Oxford: Oxford University Press.

Swenson, G. (1963) Andy Warhol, "What Is Pop Art?", interview with Gene Swenson, *Artnews* 62, November 1963.

Winograd, T. (1991) James Sheehan and Morton Sosna, eds., *The Boundaries of Humanity: Humans, Animals, Machines*, Berkeley: University of California Press.