EVERY WORD

CATARINA LEE
catarinaleecastanheira@gmail.com

Lisbon, Portugal
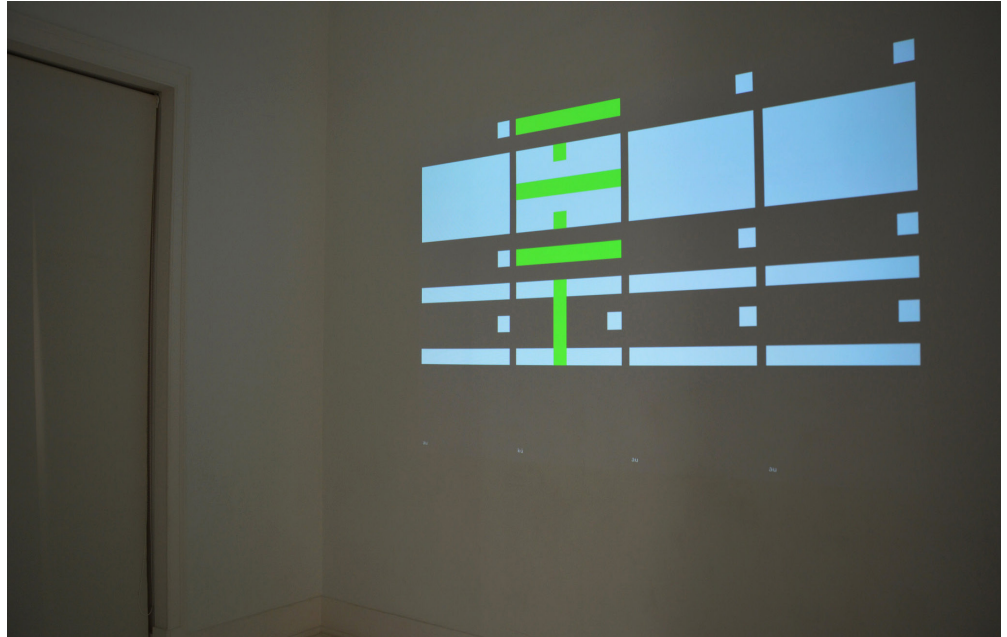
Abstract

The project *Every Word* proposes an expressive exploration of the similarities and differences between languages, through an audio-visual sequence that translates textual content into graphics and sounds. It consists of a program that is scanning through all the words of four languages simultaneously, while generating a sequence of symbols and modulating sound parameters. The aim of this installation is to explore the potential of software to translate digital textual data into new expressive forms, and eventually expose inherent dimensions and patterns in the source text. This project follows an on-going research that focus on the notion of transmutability of digital data, with a particular interest on the exploration of textual material.

Keywords

Digital Data
Text
Visualization
Sonification
Transmutability

1. TRANSMUTABILITY OF TEXTUAL DATA

The concept of transmutability, as proposed by Levin (2010) relies on the "premise that any information can be algorithmically sonified or visualized". New possibilities for linking and generating sounds and images arise, considering that, within the computer, all media objects are composed of digital code which can be regarded as raw material and translated into other forms through algorithmic manipulation (Manovich 2001, 27). This creative potential is explored through practices that rely on software as their medium, and involve articulations between the visual, auditory, and other physical or tangible realms (Levin 2010).

This work draws on a previous study and practical exploration of this notion, in which we addressed its conceptual and creative potential (Lee et al. 2014). We examined practices that explore the mutability of data in their different approaches for its reconfiguration, following either analytical or aesthetic concerns. While some projects use digital data as raw material and emphasize the translation process itself, other projects focus on real world data signals, while proposing a new expression or perception of this source material. Furthermore, depending on the nature of the source data, the process of mapping can articulate both visualization and sonification methods, namely as explored through sound visualization or image-based sonification. As such, transmutability is expressed by practices that put an emphasis on data as content, on its representation and perception, as well as on the procedures for its reconfiguration (Lee et al. 2014, 418-19). In line with this idea, we observed that text visualization and exploration has been emerging as a "growing and increasingly important subfield of information visualization" (Kucher & Kerren 2015, 117). The interest in exploring texts as source data, and the potential of finding "visual ways to make them talk" is related to the fact that "a lot of the richest information we have" is available in text formats (Viégas as cited in Heer 2010, 7) and is increasingly growing. In this sense, and considering the advances being made in text analysis (Nualart-Vilaplana et al. 2014, 224), we can identify a transformative potential worthy of exploration, which is tied to its manipulation and translation by computational means.

Therefore, more recent developments of our approach to transmutability are focused on the nature of text as source data, and the different conceptual and aesthetic intents behind its exploration. They pertain to the use of textual data *per se*, as *content* or as an *abstraction*, thus respectively exploring text as it is, reflecting on the meaning it conveys, or even emphasizing its mutability as digital data. These strategies expose the potential of translating and revealing inherent (and eventually latent or hidden) dimensions of text, relating to its *formal* specificities, *semantic* aspects, or its *abstraction*, through its mapping into a new expressive form (Lee & Ribas 2016, 212-214).

2. CONCEPT

Based on these observations, the project *Every Word* proposes a practical exploration and illustration of transmutability, focusing on the expressive potential of the nature and formal qualities of text, as well as the abstract nature of its digital encoding. We seek to explore ways of audio-visually translating textual content, in order to provide new perceptions or experiences of it through seeing and hearing. The project aims to highlight the similarities and differences between diverse languages, taking as source material a list of their vocabulary. It focuses on the

formal structure, material qualities, and internal logic of its individual elements, the words. The process departs from an analytical stance, and is gradually detached from text semantics towards a more expressive approach. By using text as raw material, and mapping textual features into graphics and sounds, the representations become ultimately abstract and emancipated from their referent. This approach seeks to conceptually emphasize the translation process, or the potential of algorithmically transforming any kind of data into a new tangible representation. In this sense, the experiences are oriented towards an aesthetic exploration of the expressive qualities of the visualizations and sonifications, and the patterns that eventually emerge from them.

**Fig. 1**
Approach scheme.
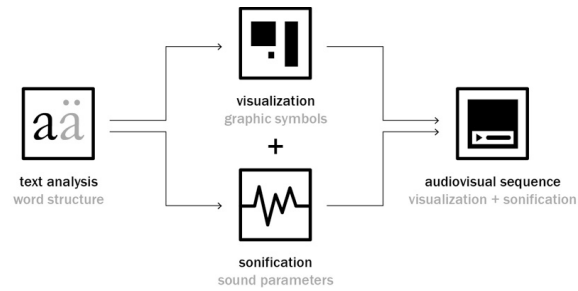
3. DEVELOPMENT

3.1. Input Data

The data source is a collection of text documents that contain a list of 'all' the existing words in different languages. Assuming that we could apply this approach to any language, in this version of the project we opted to focus on four languages: English, German, Portuguese and Italian. This selection was made because these languages use the Latin alphabetic system, but belong to different language families, being respectively two Germanic and two Romanic languages, which are prone to give rise to different expressive results. We begin with a text file containing a glossary of all the words in a selected English dictionary in alphabetical order, given that it is the most widely spoken language within the selection. We generate automatic translations of this text into German, Portuguese and Italian. The resulting lists of words are then used as a basis for the visualizations and sonifications.

3.2. Mapping Process

The mapping process is inspired by simple textual analysis and visual mapping techniques. It involves a system of correspondences between textual features, graphic symbols and sound parameters. We begin with a lexical and morphological analysis of the glossaries, extracting elements that are mapped into graphical features and used to modulate sound parameters. Rather than focusing on the syntactical structure of each language, we assume each word as an individual element, and explore its particular sequence of characters, as the basic units of the written language. Elements like the ASCII code of each character, the accentuation, or the number of characters of each word, are used to generate graphic symbols and audio frequencies.

Fig. 2
Mapping process.
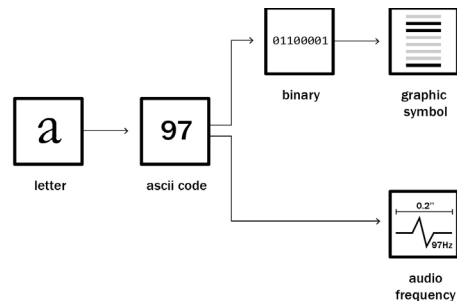
The resulting images and sounds are then displayed sequentially, character by character, and word by word, and when combined they create an audiovisual sequence. The program runs through all the words in each file, at the same time, and presents the four sequences simultaneously, in order to highlight each language's singularities.

3.3. Output

As the program is scanning through the texts, the resulting sequence presents an audiovisual reading and expression of the four languages. Since parameters such as special characters (accentuation), or particular combinations or sequences of characters (as common prefixes and suffixes) in each text are corresponded to particular symbols and sounds, visual patterns and rhythms will eventually start to emerge.

In terms of formal representation, we opted for the use of elementary figures and sounds, seeking to minimize aspects that are accessory to the audio-visual reading of the text, and taking advantage of the automatization of the computational mapping process.

Fig. 3
Snapshots (test).
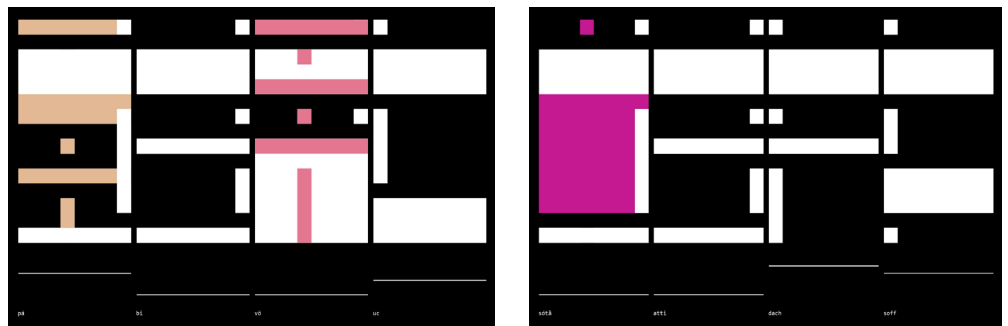
4. RESULTS

The presented work results from a closed system of correspondences between text, graphic symbols and sound parameters. The process involves analyzing and extracting elements of the text, and defining graphic symbols, as well as exploring frequency and amplitude modulation, and sound filtering techniques. When the visualizations and sonifications are combined, they are presented as a temporal sequence, ultimately becoming an abstract notation of the texts. The transient output and the closed non-variable nature of this work aims to promote a contemplative experience on the patterns and rhythms that emerge from each language.

In this manner, this work seeks to explore the creative and expressive potential of translating text into visual and auditory representations and reveal some of its hidden dimensions. As part of an on-going research, this work can be un-

derstood as an open process, where the visualizations and sonifications here described provide a starting point for further explorations, namely through the use of other sources of data and mapping processes in order to generate different dynamic results.

REFERENCES

Heer, Jeff. "A conversation with Jeff Heer, Martin Wattenberg, and Fernanda Viégas". Queue, March, v. 8, n. 3. <http://doi.acm.org/10.1145/1737923.1744741> [visited: 30-01-2017]. 2010.

Kucher, Kostiantyn; and Andreas Kerren. "Text Visualization Techniques: Taxonomy, Visual Survey, and Community Insights". Proceedings of the 8th IEEE Pacific Visualization Symposium (PacificVis'15). IEEE Computer Society Press, Hangzhou, China: 117-121, 2015.

Lee, Catarina; and Luísa Ribas. "Exploring Textual Data: Transmutability as a Creative Concept and Practice". xCoAx 2016: Proceedings of the 4th Conference on Computation, Communication, Aesthetics and X. Bergamo: 211-224, 2016.

Lee, Catarina; Luísa Ribas, and Miguel Cardoso. "On the Notion of Transmutability: From Concept to Practice". xCoAx 2014: Proceedings of the 2nd Conference on Computation, Communication, Aesthetics and X. Porto: 416-429, 2014.

Levin, Golan. "Audiovisual Software Art". *Audiovisuology: Compendium*. Daniels, Dieter and Sandra Naumann (eds). Cologne: Verlag der Buchhandlung Walther König, 271-283, 2010.

Manovich, Lev. *The Language of New Media*. Cambridge, MA: The MIT Press, 2001.

Nualart-Vilaplana, Jaume; Mario Pérez-Montoro, and Mitchell Whitelaw. "How we draw texts: a review of approaches to text visualization and exploration". El profesional de la información, May-June, v. 23, n. 3, 221-235. <http://dx.doi.org/10.3145/epi.2014.may.02> [visited: 30-01-2017]. 2014.